

VISIT...

LANZAROTE
Caliente.COM

Т Е О Р И Я АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Под редакцией д-ра техн. наук,
проф. **В.Б. Яковлева**

Допущено Министерством
образования Российской Федерации
в качестве учебника для студентов
высших учебных заведений, обучающихся
по направлению подготовки бакалавров
и магистров «Автоматизация и управление»
и направлению подготовки дипломированных
специалистов «Автоматизация и управление»



Москва
"Высшая школа"

2003

УДК 681.5
ББК 32.965
Т33

С. Е. Душин, Н. С. Зотов, Д. Х. Имаев, Н. Н. Кузьмин, В. Б. Яковлев

Рецензенты:

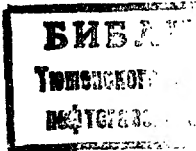
Кафедра Систем управления и информатики Санкт-Петербургского государственного института точной механики и оптики (технического университета) — зав. кафедрой — д-р техн. наук, проф. В. В. Григорьев; д-р техн. наук, проф. Г. А. Дидук (кафедра Процессов управления и информационных систем Северо-западного государственного заочного технического университета)

Т33 Теория автоматического управления: Учеб. для вузов/С. Е. Душин, Н. С. Зотов, Д. Х. Имаев и др.; Под ред. В. Б. Яковлева. — М.: Высшая школа, 2003. — 567 с.: ил.

ISBN 5-06-004096-8

В учебнике изложены основы теории управления, включающие общие сведения о системах управления и их моделях, методы анализа и синтеза линейных непрерывных и дискретных систем при детерминированных и стохастических воздействиях, методы анализа нелинейных систем, а также методы оптимизации систем управления. На базе общих принципов системного подхода рассматривается проблематика математического описания систем управления, раскрываются взаимосвязи понятий управления и информации, структуры систем и фундаментальных свойств их поведения — устойчивости, инвариантности и чувствительности. С единых позиций и с учетом применения компьютерных программ расчета и имитации изложены классические и современные методы анализа и синтеза систем управления.

Для студентов технических вузов, обучающихся по направлению «Автоматизация и управление».



УДК 681.5
ББК 32.965

ISBN 5-06-004096-8

© ФГУП «Издательство «Высшая школа», 2003

Оригинал-макет данного издания является собственностью издательства «Высшая школа», и его репродуцирование (воспроизведение) любым способом без согласия издательства запрещается.

ОГЛАВЛЕНИЕ

<i>Предисловие</i>	11
ЧАСТЬ 1. ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ И ТЕОРИИ УПРАВЛЕНИЯ ..	15
<i>Глава 1. Основные понятия теории управления</i>	15
1.1. Понятия об управлении и системах управления	15
1.2. Объекты управления	17
1.2.1. Технические объекты	18
1.2.2. Экономические объекты	18
1.2.3. Биологические системы	20
1.3. Поведение объектов и систем управления	21
1.4. Информация и принципы управления	23
1.4.1. Разомкнутые системы управления	24
1.4.2. Компенсация возмущений	25
1.4.3. Системы управления с обратной связью	26
1.4.4. Системы с компенсацией параметрических возмущений	27
1.4.5. Адаптивное управление	28
1.5. Классификация систем управления	29
1.5.1. Классификация систем управления по типу сигнала	30
1.5.2. Классификация систем автоматического регулирования по типу алгоритма	31
1.5.3. Классификация систем управления по энергетическому признаку ..	32
1.6. Задачи теории управления	34
1.7. Математические модели систем управления	35
1.7.1. Взаимодействие системы со средой	35
1.7.2. Операторы преобразования переменных	36
1.7.3. Классы моделей	38
1.7.4. Автономные системы	39
1.7.5. Модели среды	40
1.8. Способы построения моделей	40
1.8.1. Аналитический способ	40
1.8.2. Электрическая схема	42
1.8.3. Механическая поступательная система	42
1.8.4. Экспериментальный способ	44
1.9. Особенности структурных моделей систем управления	45
1.9.1. Модели систем управления с раскрытой причинно-следственной структурой	45
1.9.2. Иерархические модели систем управления	46
ЧАСТЬ 2. ТЕОРИЯ ЛИНЕЙНЫХ НЕПРЕРЫВНЫХ СИСТЕМ УПРАВЛЕНИЯ	48
<i>Глава 2. Линейные модели и характеристики систем управления</i>	48
2.1. Модели вход-выход	48

2.1.1. Дифференциальные уравнения	48
2.1.2. Передаточные функции	49
2.1.3. Временные характеристики	50
2.1.4. Частотные характеристики	51
2.1.5. Преобразование форм представления моделей вход-выход	53
2.2. Построение временных характеристик	53
2.3. Построение частотных характеристик	58
2.4. Полнота характеристик	60
2.5. Системы дифференциальных уравнений	61
2.5.1. Системы дифференциальных уравнений различных порядков	61
2.5.2. Системы дифференциальных уравнений в форме пространства состояний	62
2.5.3. Дифференциальные уравнения автономных систем	63
2.6. Построение моделей вход-выход по системе дифференциальных уравнений	64
2.6.1. Последовательное исключение переменных	64
2.6.2. Правило Крамера	65
2.6.3. Матричный способ	66
2.7. Построение моделей вход-выход по уравнениям в форме пространства состояний	67
2.7.1. Вычисление передаточной функции	67
2.7.2. Построение временных характеристик	69
2.8. Модели систем управления с раскрытой причинно-следственной структурой	72
2.8.1. Структурные схемы (С-графы)	72
2.8.2. Сигнальные графы	74
2.8.3. Причинно-следственные системы дифференциальных уравнений ..	75
2.9. Преобразование форм представления моделей с раскрытой структурой	76
2.9.1. Запись дифференциальных уравнений по графу	76
2.9.2. Построение графов по дифференциальным уравнениям	77
2.10. Характеристики систем с типовыми структурами	78
2.10.1. Характеристики последовательного соединения	79
2.10.2. Характеристики параллельного соединения	80
2.10.3. Характеристики соединения с обратной связью	81
2.11. Построение моделей вход-выход систем с раскрытой структурой	84
2.11.1. Последовательное применение правил эквивалентного преобразования графов	85
2.11.2. Формула Мэзона	86
2.12. Построение моделей с раскрытой структурой по моделям вход-выход	88
2.12.1. Построение модели в форме пространства состояний по дифференциальному уравнению n -го порядка	88
2.12.2. Построение структурных схем по передаточной функции	91
2.12.3. Типовые звенья	93
2.13. Модели систем управления высших уровней	94
2.14. Модели среды и расширенной системы	95
2.15. Неопределенность моделей систем управления	96

Глава 3. Анализ систем управления	98
3.1. Задачи анализа	98
3.2. Анализ устойчивости	99
3.2.1. Свободные движения	99
3.2.2. Условие устойчивости	100
3.2.3. Устойчивость вход-выход	101
3.3. Критерии устойчивости	101
3.3.1. Необходимое условие устойчивости	101
3.3.2. Алгебраические критерии	102
3.3.3. Частотный критерий Михайлова	104
3.4. Устойчивость систем с типовой структурой	104
3.4.1. Устойчивость систем без контуров	104
3.4.2. Устойчивость одноконтурных систем	105
3.4.3. Приближенная зависимость корней характеристического полинома одноконтурной системы от нулей и полюсов передаточной функции разомкнутой системы	107
3.5. Критерий Найквиста	110
3.6. Инвариантность систем управления	115
3.6.1. Формы инвариантности	116
3.6.2. Селективная инвариантность к степенным воздействиям	118
3.6.3. Селективная инвариантность к гармоническому воздействию	120
3.7. Инвариантность систем с типовой структурой	121
3.7.1. Последовательная компенсация воздействия	121
3.7.2. Параллельная компенсация воздействия	123
3.7.3. Инвариантность в системах с обратной связью	124
3.8. Чувствительность систем управления	127
3.9. Чувствительность систем с типовой структурой	129
3.9.1. Чувствительность системы, образованной последовательным соединением звеньев	129
3.9.2. Чувствительность системы, образованной параллельным соединением звеньев	130
3.9.3. Чувствительность системы с обратной связью	130
3.10. Чувствительность систем со сложной структурой	132
3.11. Показатели качества собственных движений систем управления	133
3.11.1. Корневые показатели качества процессов управления	135
3.11.2. Интегральные показатели качества	138
3.11.3. Частотные показатели качества	145
3.12. Показатели качества вынужденных процессов управления	146
3.13. Управляемость и наблюдаемость систем	149
3.13.1. Понятия управляемости и наблюдаемости	149
3.13.2. Критерии управляемости и наблюдаемости	153
3.13.3. Принцип дуальности	157
Глава 4. Синтез систем управления	158
4.1. О синтезе систем управления	158
4.2. Задачи синтеза систем управления	161
4.3. Стабилизация неустойчивых объектов	166
4.3.1. Размещение корней характеристического полинома. Операторный метод	167
4.3.2. Размещение собственных значений матрицы дифференциальных уравнений в форме пространства состояний	171
4.3.3. Аналитическое конструирование регуляторов	175
4.3.4. Синтез наблюдателя состояний	177

4.4. Синтез систем управления, инвариантных к возмущениям	178
4.4.1. Синтез статических систем	179
4.4.2. Синтез астатических систем по требованию к точности подавления степенных возмущений	181
4.4.3. Синтез систем по требованию к точности подавления гармонических возмущений	183
4.5. Синтез следящих систем из условия требуемой точности воспроизведения	184
4.6. Коррекция систем управления	187
4.7. Методики формирования желаемых передаточных функций	192
4.8. Вычисление передаточных функций корректирующих устройств	197
4.8.1. Последовательная коррекция	199
4.8.2. Местная обратная связь	202
4.9. Параметрический синтез систем автоматического управления ..	206

ЧАСТЬ 3. ТЕОРИЯ ЛИНЕЙНЫХ ДИСКРЕТНЫХ СИСТЕМ УПРАВЛЕНИЯ 212

Глава 5. Общие сведения о дискретных системах автоматического управления 212

5.1. Классификация дискретных систем по виду квантования	212
5.2. Понятия об импульсных системах автоматического управления	214
5.3. Обобщенные структурные схемы импульсных автоматических систем	217

Глава 6. Модели линейных дискретных систем управления 222

6.1. Решетчатые функции и разностные уравнения	222
6.2. Математическое описание идеального импульсного элемента ...	227
6.3. Уравнения и импульсная передаточная функция разомкнутой импульсной системы	236
6.4. Частотные характеристики импульсных систем	243
6.5. Логарифмические частотные характеристики импульсных систем	247
6.6. Представление дискретных систем в пространстве состояний ...	253

Глава 7. Анализ импульсных систем автоматического управления ... 260

7.1. Структурные схемы и передаточные функции замкнутых импульсных систем	260
7.2. Процессы в импульсных системах	264
7.3. Оценка точности импульсных систем автоматического управления в установившемся режиме	268
7.4. Устойчивость импульсных систем	271

Глава 8. Синтез импульсных систем управления 273

8.1. О синтезе импульсных систем	273
8.2. Построение желаемых частотных характеристик	276
8.3. Синтез непрерывных корректирующих устройств	282
8.4. Синтез дискретных корректирующих устройств	286

ЧАСТЬ 4. НЕЛИНЕЙНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ 292

Глава 9. Нелинейные модели систем управления 292

9.1. Необходимость в нелинейных моделях	292
9.2. Безынерционные нелинейные элементы	295

9.3. Динамические нелинейные элементы	299
9.4. Нелинейные модели с раскрытой структурой	300
9.5. Расчетные формы нелинейных моделей	301
Глава 10. Анализ равновесных режимов	304
10.1. Равновесные режимы	304
10.2. Определение равновесных режимов по дифференциальным уравнениям	305
10.3. Определение равновесных режимов и статических характеристик систем с типовой структурой	307
10.4. Определение статических характеристик систем с произвольной структурой	310
10.5. Равновесные режимы в системах управления и катастрофы	312
Глава 11. Анализ поведения систем на фазовой плоскости	316
11.1. Метод фазовой плоскости	316
11.2. Поведение нелинейных систем в окрестности положений равновесия	320
11.3. Построение фазовых портретов нелинейных систем	324
11.3.1. Метод изоклии	324
11.3.2. Примеры аналитического построения фазовых портретов	325
11.4. Связь фазовых траекторий со временем	328
11.5. Особенности фазовых портретов нелинейных систем	330
Глава 12. Устойчивость положений равновесия	333
12.1. Понятие об устойчивости невозмущенного движения	333
12.2. Первый метод Ляпунова	335
12.2.1. Применение первого метода Ляпунова к дифференциальным уравнениям в форме Коши	335
12.2.2. Исследование устойчивости по дифференциальным уравнениям n -го порядка	336
12.2.3. Исследование устойчивости по моделям в форме структурной схемы	338
12.3. Второй (прямой) метод Ляпунова	341
12.4. Применение второго метода Ляпунова	344
12.5. Частотный метод исследования абсолютной устойчивости	346
Глава 13. Исследование периодических режимов методом гармонического баланса	350
13.1. Основные положения метода гармонического баланса	350
13.2. Гармоническая линеаризация нелинейного элемента	353
13.3. Способы вычисления коэффициентов гармонической линеаризации	354
13.3.1. Вычисление коэффициентов гармонической линеаризации усреднением по фазе	354
13.3.2. Вычисление коэффициентов гармонической линеаризации усреднением по множеству значений входа нелинейного элемента	355
13.3.3. Гармоническая линеаризация соединений нелинейных элементов	358

13.4. Определение параметров периодических режимов	359
13.4.1. Метод Л.С.Гольдфарба	359
13.4.2. Метод А.А.Вавилова	361
13.4.3. Метод Е.П.Попова	363
13.5. Устойчивость периодических режимов	365
13.6. Чувствительность периодических режимов	367

ЧАСТЬ 5. СИСТЕМЫ УПРАВЛЕНИЯ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ 370

Глава 14. Общие сведения о случайных воздействиях и процессах в системах управления 370

14.1. Случайные воздействия	370
14.2. Случайный процесс и его основные характеристики	371
14.3. Стационарные случайные процессы	376
14.3.1. Определения стационарных случайных процессов	376
14.3.2. Эргодические случайные процессы	378
14.3.3. Основные свойства корреляционной функции	382
14.3.4. Спектральная плотность и ее связь с коррекционной функцией	382
14.4. Модели случайных воздействий	385

Глава 15. Анализ линейных систем при случайных воздействиях 389

15.1. Преобразование случайного сигнала линейным звеном	389
15.2. Объект управления при случайных воздействиях	393
15.3. Замкнутая система при случайных воздействиях	395
15.4. Способы вычисления дисперсии случайного сигнала	398
15.4.1. Аналитический способ	398
15.4.2. Графоаналитический способ	404
15.5. Применение идеального белого шума в качестве модели внешней среды	405
15.6. Вычисление дисперсии ошибки в системах различного порядка	409
15.6.1. Оценка точности преобразования стационарных случайных сигналов в установившихся режимах	409
15.6.2. Вычисление дисперсии ошибки в системах первого порядка	411
15.6.3. Вычисление дисперсии ошибки в системах второго порядка	413
15.6.4. Вычисление дисперсии ошибки в системах третьего порядка	418

Глава 16. Синтез линейных систем при случайных воздействиях 421

16.1. Постановка задач синтеза	421
16.2. Синтез оптимальной системы с заданной структурой	424
16.3. Синтез оптимальной системы с произвольной структурой	428
16.3.1. Описание оптимальной системы интегральным уравнением Винера-Хопфа	428
16.3.2. Определение оптимальной передаточной функции с учетом физической реализуемости системы (фильтр Винера)	432
16.3.3. Определение передаточной функции системы с ограниченной длительностью переходного процесса (фильтр с конечной памятью)	437
16.4. Синтез оптимальной системы в пространстве состояний (фильтр Калмана-Бьюси)	441

Глава 17. Нелинейные системы при случайных воздействиях	443
17.1. Преобразование случайного сигнала нелинейным звеном	443
17.2. Статистическая линеаризация нелинейного звена	446
17.3. Исследование нелинейных систем методом статистической линеаризации	450
ЧАСТЬ 6. ОПТИМАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ	459
Глава 18. Условия оптимальности процессов в динамических системах	459
18.1. Принцип максимума	459
18.1.1. Постановка задачи	459
18.1.2. Формулировка принципа максимума	461
18.1.3. Применение принципа максимума	463
18.1.4. Условия трансверсальности	464
18.1.5. Задачи Больца и Майера	466
18.2. Численные методы расчета оптимального управления на основе принципа максимума	467
18.2.1. Решение краевой задачи путем поиска начального значения сопряженного вектора	468
18.2.2. Приращение функционала	468
18.2.3. Метод градиента для решения задачи Майера	469
18.2.4. Метод последовательных приближений	471
18.2.5. Преобразование к задаче нелинейного программирования	472
18.3. Метод динамического программирования	473
18.3.1. Принцип оптимальности	473
18.3.2. Дискретные детерминированные системы	474
18.3.3. Непрерывные детерминированные системы	476
18.3.4. Непрерывные стохастические системы	479
Глава 19. Системы, оптимальные по быстродействию	480
19.1. Определение алгоритма управления	480
19.1.1. Вырожденность управления	483
19.1.2. Теорема об n -интервалах	483
19.2. Расчет моментов переключения	486
19.2.1. Описание объекта в пространстве состояний	486
19.2.2. Описание объекта в виде уравнения вход-выход	486
19.2.3. Численный метод расчета на ЭВМ оптимальных по быстродействию процессов	489
19.3. Замкнутые оптимальные по быстродействию системы	491
19.4. Синтез замкнутой оптимальной по быстродействию системы второго порядка	492
19.4.1. Учет изменения параметров объекта	495
19.5. Оптимальные по быстродействию системы с ограниченными ресурсами	498
19.5.1. Оптимизация по критерию расхода топлива	499
19.5.2. Оптимизация по расходу энергии	503
19.5.3. Численный алгоритм оптимизации по расходу топлива и энергии для линейного объекта управления	507
Глава 20. Оптимизация динамических систем по квадратичному критерию	511
20.1. Постановка задачи	511
20.2. Детерминированный оптимальный регулятор с фиксированным временем перехода при полной информации о состоянии объекта	515

20.2.1. Уравнение Риккати	518
20.2.2. Свойства решения уравнения Риккати	520
20.3. Детерминированный регулятор с бесконечным временем перехода при полной информации о состоянии	520
20.3.1. Устойчивость замкнутой системы с оптимальным регулятором ..	522
20.3.2. Учет постоянно действующих возмущений	523
20.3.3. Вычисление матрицы стационарного оптимального регулятора ..	524
20.3.4. Асимптотическое распределение корней	528
20.3.5. Оптимальность в частотной области	528
20.3.6. Критерий обобщенной работы	529
20.4. Стохастический непрерывный оптимальный регулятор при полной информации о состоянии	530
20.5. Дискретные системы, оптимальные по квадратичному критерию	532
20.5.1. Дискретные детерминированные оптимальные системы	533
20.5.2. Оптимальные стохастические дискретные системы	535

Глава 21. Оптимальные системы при неполном измерении вектора состояния

21.1. Оптимальная система с детерминированным наблюдателем полного порядка	537
21.2. Детерминированный наблюдатель пониженного порядка	540
21.3. Оптимальный стохастический непрерывный наблюдатель	543
21.3.1. Стационарный оптимальный наблюдатель	546
21.3.2. Сингулярная (вырожденная) задача наблюдения	548
21.4. Стохастическая непрерывная оптимальная система управления при неполной информации. Принцип разделения	551
21.5. Оптимальный дискретный наблюдатель	553
21.5.1. Стохастическая оптимальная дискретная система управления при неполном измерении состояния. Принцип разделения	556

Заключение

Список литературы

ПРЕДИСЛОВИЕ

Окружающий нас мир представляет собой совокупность взаимосвязанных социальных, экономических, технических, биологических и иных систем, процесс функционирования каждой из которых включает получение информации и ее анализ, принятие решений и их исполнение, т. е. выполнение основных операций любого процесса управления. В наиболее общей трактовке это — организация целенаправленного взаимодействия энергии, вещества и информации. Приоритетной и конечной целью использования информации является управление, т. е. выработка и реализация управляющих действий. Универсальность принципов управления позволяет применять их к объектам любой природы, в том числе к самим процессам получения, передачи и обработки информации. Таким образом, и сами процессы управления могут рассматриваться как объекты управления систем более высоких уровней иерархии.

Дисциплина «Теория автоматического управления» или «Теория управления» формирует у будущих инженеров и исследователей системное мышление и целостное видение явлений мира техники, природы, социальной среды, синтезирует знания студентов по математике, физике, другим естественным наукам, дает наглядные примеры их конструктивного применения.

В последние годы практика проектирования и реализации алгоритмов управления претерпела значительные изменения, обусловленные интенсификацией производственных технологий, актуализацией проблем экологии, революционными достижениями в области получения, передачи и обработки информации. Применение непрерывно развивающихся компьютерных программ моделирования и имитации, анализа и синтеза систем управления в большинстве случаев устраняет вычислительные проблемы, связанные с исследованиями и разработками таких систем. Вместе с тем накопленный опыт позволяет утверждать, что расширение области применения теории управления, т. е. круга решаемых на ее основе задач, внедрение новых технологий исследования и

проектирования систем управления относительно мало отразились на основах названной учебной дисциплины.

По основам теории автоматического управления в нашей стране издан целый ряд учебников и учебных пособий. Однако они написаны относительно давно, и сохранившихся экземпляров в настоящее время недостаточно для обеспечения учебного процесса в технических вузах. Это связано с тем, что дисциплины «Теория автоматического управления» и «Теория управления» входят в состав федеральной компоненты Государственного образовательного стандарта многих специальностей и направлений высшего профессионального образования.

Вполне очевидна сложность задачи написания нового учебника. Особую трудность здесь представляет отбор материала, относящегося к «основам», из всего того необъятного, что входит в «классическую» и «современную» теорию управления. По этому вопросу отсутствует единая точка зрения, поэтому право на существование имеют самые разные мнения, представляющие различные школы.

Традиционное изложение основ теории управления обычно требует большого объема учебника, поскольку значительное место здесь уделяется вспомогательному материалу: математическим основам, способам, приемам и методикам упрощения вычислений, подробному описанию свойств типовых динамических звеньев и т. п. Изложение тех же основ в ряде книг по «современной» теории управления базируется на абстрактном, преимущественно формально-алгебраическом подходе, трудном для восприятия студентами при изучении фундаментальных свойств и раскрытии «механизмов» управления. В ряде учебников по основам теории автоматического управления материал систематизируется не по задачам теории управления, а по формам представления математических моделей (передаточные функции, структурные схемы или системы дифференциальных уравнений в матричной форме пространства состояний) и по методам расчета (частотные, графо-аналитические или временные, аналитические), что отражает не столько суть, сколько — этапы развития дисциплины.

В предлагаемом учебнике авторами сделана попытка изложения основ теории управления без ее разделения на «классическую» и «современную». Другой его особенностью является четкое выделение материалов, посвященных математическим моделям, анализу и синтезу систем управления.

Учебник содержит сравнительно большой объем материала по математическим моделям. Это объясняется тем, что именно особенности математических моделей в значительной мере определяют рамки любой теоретической дисциплины. Характерной чертой структурных моделей, используемых в теории управления, отличающих их от динами-

ческих (поведенческих) моделей, рассматриваемых в других учебных дисциплинах, является причинно-следственный характер взаимодействий элементов систем, что отражает информационно-алгоритмический подход, принятый в теории управления и теории связи (коммуникаций). Еще одной особенностью предлагаемого учебника является попытка систематизации материала по задачам, а не по методам их решения. Авторы стремились к сбалансированному рассмотрению фундаментальных свойств систем управления — устойчивости, инвариантности, чувствительности. В книге усилено внимание к теоретико-системному аспекту проблематики моделирования и анализа систем управления, объяснению свойств поведения систем в зависимости от характеристик образующих их частей и особенностей структуры причинно-следственных связей между ними. Сознательно сокращен объем материала по анализу устойчивости линейных систем, так как применение ЭВМ снижает роль некоторых традиционных методов констатации устойчивости систем этого класса. По той же причине не приводятся различные графические процедуры и номограммы, используемые при анализе и расчете показателей качества типовых систем, формировании желаемых характеристик и т. д. Отсутствует также подробное описание типовых динамических звеньев — их, по мнению авторов, следует рассматривать как примеры при проведении лабораторных практикумов.

В книгу включены разделы, посвященные методам оптимизации систем управления, которые находят широкое применение в инженерной практике. Хотя эта часть материала выходит за рамки обычного курса теории управления, авторы сочли необходимым дополнить учебник соответствующими разделами.

Книга состоит из шести частей. Первая часть посвящена краткому неформальному изложению основных понятий теории управления. Во второй части приводятся формы представления линейных непрерывных конечномерных моделей, методы анализа устойчивости, инвариантности и чувствительности, а также задачи и методы синтеза систем управления. В третьей части рассматриваются линейные дискретные и импульсные системы управления, приводятся общие сведения об автоматических системах, анализ и синтез которых целесообразно проводить по дискретным моделям, даются методы анализа и синтеза импульсных систем. В четвертой части книги вводятся нелинейные модели систем управления, обосновывается их необходимость, даются методы анализа равновесных режимов, анализа поведения динамических систем на фазовой плоскости, методы исследования устойчивости положения равновесия, а также приближенный метод исследования периодических режимов в нелинейных системах с обратной связью. Пятая

часть книги посвящена анализу и синтезу систем управления при случайных воздействиях. Содержание этой части в значительной степени базируется на материалах лекций, которые в течение многих лет читал доцент кафедры Автоматики и процессов управления (АПУ) Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В. И. Ульянова (Ленина) (СПбГЭТУ) Э. В. Сергеев. В шестой части учебника изложены методы оптимизации управлений и алгоритмов управления. Ряд задач теории управления, рассмотренных в предыдущих разделах, здесь приводится в более общей постановке, даются строгие методы их решения. В отличие от предыдущих частей книги, где задачи и методы теории управления излагались на примерах одномерных систем, здесь рассматриваются многомерные системы управления.

Авторы книги — преподаватели кафедры АПУ СПбГЭТУ «ЛЭТИ» — стремились отразить лучшие традиции научно-педагогической школы кафедры в области теории управления. В основе учебника лежат курсы лекций по теории управления, прочитанные в разные годы профессорами В. Б. Яковлевым, В. А. Олейниковым, Д. Х. Имаевым, А. И. Солодовниковым, В. А. Тереховым, доцентами Н. С. Зотовым, Ф. Ф. Котченко, Н. Н. Кузьминым и Э. В. Сергеевым. На изложение материала сильное влияние оказали сформулированные членом-корреспондентом АН СССР А. А. Вавиловым принципы системного подхода к построению моделей, анализу и эволюционному синтезу систем управления.

Большую помощь при обсуждении содержания учебника оказали профессора А. И. Солодовников, В. А. Терехов и А. Э. Янчевский, доценты А. А. Алексеев, В. Г. Григорян и Л. Б. Пошехонов. Подготовка рукописи выполнена инженерами Т. Р. Белинской, А. Н. Ветровым и Т. О. Михайловой. Всем им авторы выражают свою искреннюю благодарность.

Авторы весьма признательны коллективу кафедры Систем управления и информатики С.-Петербургского государственного института точной механики и оптики (технического университета), а также профессору Г. А. Дидуку за рецензирование рукописи и весьма ценные замечания, которые были учтены при ее окончательном редактировании.

Авторы настоящего издания понимают, что написание учебника по основам теории управления — дело весьма непростое и ответственное, поскольку имеют право на существование различные взгляды как на содержание такого учебника, так и на методику преподавания соответствующей учебной дисциплины. Авторы с благодарностью примут любые замечания и предложения коллег.

Авторы

ОБЩИЕ СВЕДЕНИЯ О СИСТЕМАХ И ТЕОРИИ УПРАВЛЕНИЯ

Глава 1 ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ УПРАВЛЕНИЯ

1.1. ПОНЯТИЯ ОБ УПРАВЛЕНИИ И СИСТЕМАХ УПРАВЛЕНИЯ

Под *управлением* понимается совокупность операций по организации некоторого процесса для достижения определенных целей.

Рассмотрим пример процесса судовождения (рис. 1.1). Целью управления является поддержание курса движущегося судна ψ равным заданному ψ_3 . Говорят, что судно является *объектом управления* (ОУ). *Возмущающие воздействия* f — ветер, волны, течения — приводят к отклонению курса от заданного $\Delta\psi = \psi_3 - \psi$. Для ослабления влияния возмущений и ликвидации отклонения используется управляющий орган — руль. Изменение положения руля μ является *управляющим воздействием* — завершающей операцией по организации процесса судовождения для достижения заданного курса.

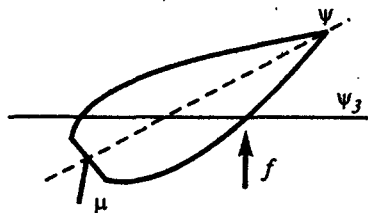


Рис. 1.1. Пример процесса управления

Анализируя пример управления судном, можно выделить следующие операции:

- 1 — получение информации о цели (заданном курсе);
- 2 — получение информации о состоянии процесса и среды;
- 3 — выявление соответствия текущего состояния процесса цели управления и принятие решения об оказании управляющего воздействия

на объект (можно предположить, что полученная информация обрабатывается по некоторому *алгоритму*);

4 — исполнение принятого решения.

Первые две операции отражают информационный, третья — алгоритмический, а четвертая — энергетический аспекты управления.

Если все операции управления осуществляются человеком, то это *ручное* управление. Судно крупного тоннажа обычно снабжено приводом руля (рулевой машиной), т. е. человек освобожден от физических усилий. В подобных случаях говорят о *механизированном* управлении. Для этого необходимы исполнительный механизм и источник энергии. *Автоматическое* управление осуществляется без непосредственного участия человека. При этом не только энергетические, но также информационные и алгоритмические операции управления передаются автоматам.

Операции автоматического управления реализуются в следующих основных функциональных элементах: *объекте управления, измерительных элементах, управляющем устройстве и исполнительном механизме.*

Совокупность взаимосвязанных и взаимодействующих элементов образует *систему управления*, если она способна к достижению цели. Системы управления представляют собой класс динамических систем, отличающихся наличием цели и специальной организации, позволяющей достичь эту цель. Подобные системы функционируют в живой и неживой природе, в самых различных областях техники, экономики, биологии. Однако, независимо от области функционирования, их объединяет в особый класс ряд системных свойств и признаков.

На рис. 1.2 условно изображена система управления (СУ) курсом судна. Объектом управления является движущееся судно; измерительным элементом — гирокопический компас (ГК), ось которого направлена вдоль заданного курса; управляющим устройством — авторулевой; исполнительным механизмом — привод руля. Линии со стрелками указывают направление причинно-следственных связей элементов. Предполагается, что элементы СУ обладают свойством однонаправленности. Например, изменение курса судна приводит к изменению сигнала на выходе гирокопического компаса, но изменение этого сигнала, очевидно, не может изменить курс судна. Изменение положения руля приводит к изменению курса движущегося судна, но изменение курса не может изменять положение руля. Свойство однонаправленности элементов, а также причинно-следственный характер их взаимодействия являются принципиальной особенностью подхода, принятого в теории управления.



Рис. 1.2. Система управления курсом судна

Схема автоматической системы управления (см. рис. 1.2) является первым уровнем абстракции, необходимой для последующего теоретического анализа. На схеме выделены и упрощенно изображены конкретные элементы системы — судно, привод руля, гирокомпас и др. Теория управления изучает закономерности управления, присущие объектам любой природы, поэтому имеет дело с более абстрактными описаниями.

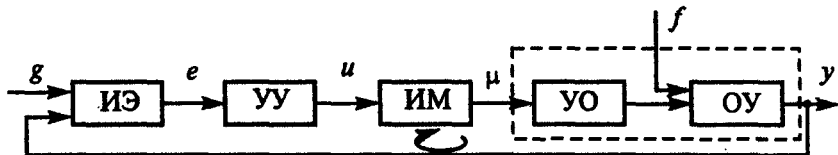


Рис. 1.3. Структура системы управления с обратной связью

Следующий уровень абстракции иллюстрирует рис. 1.3, где приведена структура системы с указанием функций элементов: объект управления (ОУ); управляющий орган (УО), обычно конструктивно входящий в объект; исполнительный механизм (ИМ); управляющее устройство (УУ); измерительный элемент (ИЭ). Там же даны и обозначения для основных переменных систем управления: y — управляемая переменная (выход объекта и системы); μ — управляющее воздействие (энергетическое); u — управляющее воздействие (информационное); e — отклонение управляемой переменной от заданного значения g или ошибка системы; f — возмущающее воздействие среды. Схема на рис. 1.3 является достаточно общей и отражает существенные с позиций теории управления черты систем управления с обратной связью.

1.2. ОБЪЕКТЫ УПРАВЛЕНИЯ

Объект управления (иначе, управляемый объект) реализует процесс, который необходимо организовать для достижения поставленных целей.

В теории управления объект рассматривается как преобразователь переменных, входящих в него, в переменную выхода $y(t)$, как это показано

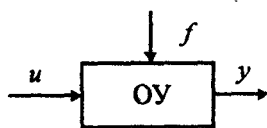


Рис. 1.4. Абстрактный объект управления

на рис. 1.4. Цель управления, в первую очередь, определяет ограничения на переменную выхода объекта $y(t)$. Неконтролируемые воздействия среды $f(t)$, называемые *возмущениями*, вызывают нежелательные отклонения выхода объекта. Для уменьшения этих отклонений на объект оказывают соответствующие *управляющие* воздействия $u(t)$.

Если переменные входа и выхода объекта являются векторами, то объект называют *многомерным*. Например, задача управления подводной лодкой или самолетом (воздушным судном) связана с рассмотрением в качестве управляемых переменных не только курсового угла ψ , но также углов крена и тангажа.

Теория управления изучает общие закономерности, присущие системам управления независимо от их природы. Объекты управления могут быть техническими, экономическими, биологическими, социальными, военными и др.

1.2.1. Технические объекты

Примерами технических объектов являются движущиеся объекты (суда, самолеты, автомобили), объекты энергетики (печи, теплообменники, парогенераторы, электрические генераторы, двигатели и другие преобразователи энергии), объекты химической промышленности (реакторы, автоклавы, мешалки и другие аппараты), объекты машиностроения и металлургии (станки и обрабатывающие центры, прокатные станы), бытовая техника (утюги, холодильники, кондиционеры, устройства бытовой электроники) и т. д. Управляемыми переменными для технических объектов являются физико-химические переменные, характеризующие их состояние или поведение (механические координаты, температура, электрическое напряжение, давление, концентрация компонентов в смесях и т. д.). Теория автоматического управления долгое время развивалась применительно именно к техническим системам, объекты которых наиболее изучены — во многих случаях для них имеются достаточно адекватные математические модели или возможность их аналитического построения [3, 13, 25, 38, 42, 64, 72, 80].

1.2.2. Экономические объекты

Экономическими объектами являются: цех, завод, отрасль, предприятия торговли и сферы услуг, банки, страховые компании и т. п. Общим для них является то, что в качестве одной из обобщенных управляемых переменных рассматривается прибыль.

Приведем простейшие примеры таких объектов и соответствующих задач (целей) управления [14, 23, 99].

Пример 1. Задача определения оптимального объема заказа. Фирма потребляет некоторую продукцию в одном и том же объеме M единиц в неделю и должна регулярно возобновлять заказы для пополнения запасов. Каждый очередной заказ независимо от его объема сопряжен с некоторыми накладными расходами K по размещению заказа и его реализации. С другой стороны, с хранением приобретаемой продукции связаны расходы на складирование, страхование и содержание складских помещений. Допустим, что за хранение единицы потребляемой продукции в течение недели взимается плата в размере h . Обозначим через Q объем заказа. Управляемой переменной в таком объекте является среднее значение суммарных затрат в расчете на одну неделю S_3 . Накладные расходы K и расходы на хранение h принимаются за возмущения. Целью управления является минимизация значения управляемой переменной

$$S_3(Q, K, h) \rightarrow \min_Q$$

Пример 2. Динамическая задача управления запасами. Пусть необходимо разработать календарную программу выпуска некоторого изделия на плановый период, состоящий из N отрезков. Предполагается, что для каждого из этих отрезков имеется точный прогноз спроса на выпускаемую продукцию, причем для разных отрезков спрос неодинаков. На экономические показатели производства влияют объемы изготавливаемых партий, поэтому фирме нередко бывает выгодно выпустить на каком-то отрезке времени продукцию в объеме, превышающем спрос в пределах этого отрезка, и хранить излишки, используя их для удовлетворения последующего спроса. Задачей управления является реализация такой программы выпуска, при которой общая сумма затрат на производство и содержание запасов минимизируется при условии полного и своевременного удовлетворения спроса на продукцию. Здесь оптимальное управление ищется в виде функции дискретного времени.

Пример 3. Задача планирования производства. Фирма выпускает два типа тяжелого оборудования — землеройные машины (E) и машины для лесоразработок (F). Прогнозирование спроса показало, что можно продать столько машин типа E и F , сколько фирма может произвести. Фирма получает прибыль \$5000 от продажи каждой машины типа E и \$4000 — машины типа F . Оборудование производится в двух филиалах фирмы A и B , каждый из которых располагает бюджетом времени в 100 ч. Производство одной машины E требует 10 ч в филиале A и 20 ч — в филиале B . Производство одной машины F требует 16 ч в филиале A и 10 ч — в филиале B . Для сохранения позиций фирмы на рынке необходимо производить не менее одной машины F на каждые три машины E . Основной потребитель заказал в сумме не менее 5 машин в любой комбинации. Объект управления в данном примере имеет два управляющих воздействия — объемы выпуска машин типа E и типа F . Управляемая переменная — прибыль P должна быть максимальной

$$P(E, F) \rightarrow \max$$

при удовлетворении ограничений. Это цель управления.

Разнообразные по характеру своей деятельности организации производственной и непроизводственной сфер, представляют собой сложные объекты, экономическая эффективность функционирования которых существенно зависит от качества организационного управления. Научной дисциплиной, занимающейся решением задач такого управления, является *исследование операций* [23]. Методами исследования операций решаются задачи управления деятельностью той или иной организации, в частности, задачи календарного планирования производства и управления запасами, задачи комплектования штатов и др. Сюда же относятся определение ассортимента выпускаемой продукции, разработка долгосрочных программ расширения производства, проектирования сети складских помещений в системе оптовой торговли и т.д.

Управление экономическими объектами является задачей, вообще говоря, более сложной, чем управление техническими объектами. Одна из причин заключается в том, что в современной экономике производственно-технические, конъюнктурно-коммерческие и прочие факторы находятся в сложной взаимной зависимости, а внешние экономические факторы содержат элементы неопределенности. Принятие управленческих решений, которые были бы одновременно и реальными, и экономически выгодными, является достаточно сложной задачей.

1.2.3. Биологические системы

Живые системы — от простейших клеток до сложных организмов — поддерживают свою жизнедеятельность благодаря действию в них механизма управления. Процессы управления в живом организме регулируют ритм сердца и дыхания, температуру тела и т. д. В живой природе механизмы управления настолько органично встроены в объекты, что не всегда можно выделить такие функциональные элементы, присущие техническим системам, как объект управления, измерительные элементы, управляющие устройства и исполнительные механизмы. Однако и здесь методы теории управления помогают исследователю глубже понять процессы [24, 48].

Для демографов, экологов, эпидемиологов, микробиологов и других специалистов большой практический интерес представляют модели управления, объясняющие механизмы изменения численности биологических популяций (саморегуляция). Такие модели могут оказаться полезными и для принятия решений в области биологического управления, например, обосновать целесообразность использования при борьбе с насекомыми-вредителями ядохимикатов или бе-

запасных мер биологического характера (в частности, естественных врагов этих вредителей).

1.3. ПОВЕДЕНИЕ ОБЪЕКТОВ И СИСТЕМ УПРАВЛЕНИЯ

В теории управления, прежде всего, интересуются *поведением* или динамическими свойствами элементов и систем — изменением во времени состояния объекта, других существенных переменных, иными словами, *процессами управления*.

Рассмотрим поведение судна как объекта управления (рис. 1.5, а). Переменные системы управления (см. рис. 1.2): курс судна — угол $\psi(t)$; электрическое напряжение на выходе гироскопического компаса, пропорциональное отклонению курса $\Delta\psi(t)$; угловое положение руля $\mu(t)$ и другие являются функциями времени t , характеризующими поведение системы. Эти переменные, имеющие различную физическую природу, называют *сигналами*. Сигналы служат носителями информации, а какой-либо их изменяющийся параметр (часто — уровень) кодирует эту информацию.

Пусть в момент времени, принятый за начальный, резко изменилось направление и сила ветра, воздействие которого $f(t)$ на рис. 1.5, б условно изображено в виде функции времени. При закреплённом руле курс судна будет изменяться (рис. 1.5, в). После исчезновения возмущения судно не возвращается на исходный курс, но и более не отдаляется от него. Говорят, что это объект без *самовыравнивания*. С точки зрения специалиста по управлению судно является нейтральным в смысле устойчивости объектом, накапливающим (интегрирующим) действие возмущений.

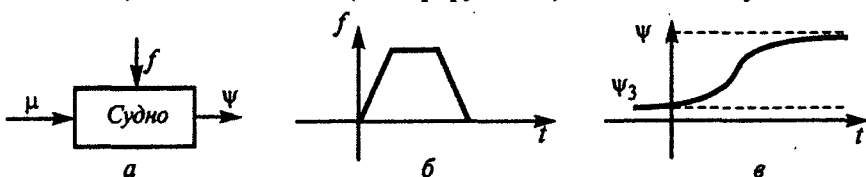


Рис. 1.5. Поведение неуправляемого судна

Рассмотрим **тепловой объект** (электрическую печь, уют и т. п.). Пусть в момент времени, принятый за начальный, на объект действует возмущение в виде кратковременного отключения напряжения (рис. 1.6, а). Температура объекта начинает понижаться (рис. 1.6, б). После исчезновения возмущения — повторного включения напряжения — температура постепенно повышается и стремится к прежнему значению. Это пример **устойчивого объекта**, состояние которого возвращается к

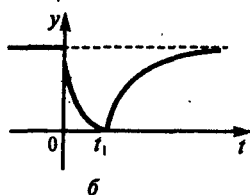
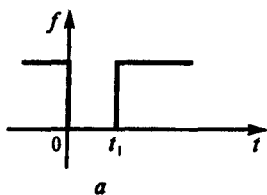


Рис. 1.6. Поведение теплового объекта

исходному после снятия возмущения. Говорят также, что объект обладает самовывравниванием.

Рассмотрим пример **механического объекта** — маятника (рис. 1.7, а). Если возмущение привело к отклонению маятника от верхнего положения равновесия, то после исчезновения возмущения маятник не только не возвращается в исходное состояние, а наоборот, отдаляется от него и начинает совершать колебания относительно нижнего положения равновесия. Верхнее положение маятника является примером неустойчивого состояния (рис. 1.7, б). Отметим, что нижнее положение маятника нейтрально в смысле устойчивости — это так называемая колебательная граница устойчивости.

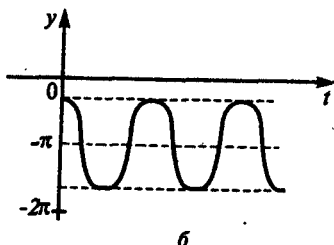
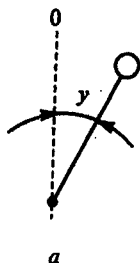


Рис. 1.7. Поведение маятника

Свойство поведения, когда после исчезновения возмущения состояние объекта стремится к исходному, называют **устойчивостью**. Устойчивость объектов и систем управления является фундаментальным свойством, необходимым для их нормального функционирования. Целью создания многих систем управления является стабилизация неустойчивых состояний (режимов) объектов.

Вернемся к примеру, проиллюстрированному рис. 1.1. Теперь положим, что судно управляется опытным рулевым. Поведение объекта

изменится качественным образом — управляемое судно после завершения переходного процесса возвращается на исходный курс (рис. 1.8), т. е. система управления, образованная судном и рулевым, является устойчивой.

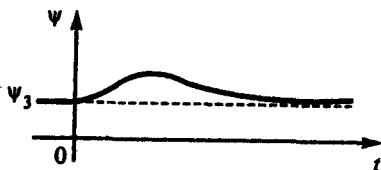


Рис. 1.8. Поведение управляемого судна

Помимо устойчивости, важным свойством, характеризующим системы управления, является *инвариантность* — независимость (слабая зависимость) управляемой переменной от действующих возмущений. Многие объекты имеют устойчивый режим функционирования, однако действие возмущений вызывает недопустимые отклонения управляемых переменных. Тогда создаются системы управления, основной целью которых является *компенсация* действия этих возмущений или их ослабление.

Поведение системы будет иным, если изменятся свойства ее элементов. Однако изменения свойств объекта или других элементов системы не должны приводить к существенным изменениям характера процессов и ни в коем случае — к потере устойчивости. Система должна быть *грубой* и малочувствительной или, как еще говорят, *робастной*.

1.4. ИНФОРМАЦИЯ И ПРИНЦИПЫ УПРАВЛЕНИЯ

Управление и информация — понятия, неразрывно связанные между собой. Управление невозможно без достаточной информации о:

- цели управления;
- возмущениях среды;
- состоянии объекта управления;
- характеристиках объекта.

В зависимости от характера и полноты доступной информации реализуют различные *принципы управления*.

1.4.1. Разомкнутые системы управления

Пусть имеется полная априорная информация о цели управления, известны характеристики объекта, а возмущающие воздействия отсутствуют. Если цель управления задана в виде

$$y = y^* = \text{const},$$

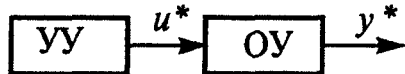


Рис. 1.9. Разомкнутая система управления

т. е. необходимо поддержание режима, заданного технологом или специалистом по эксплуатации объекта, то требуемый постоянный уровень управляющего воздействия u^* легко определяется по статической характеристике

объекта. Если этот уровень допустим, а режим объекта устойчив, то цель управления может быть достигнута в системе, реализующей принцип разомкнутого управления (рис. 1.9).

Для многих производственных объектов оптимальному режиму, который должен поддерживаться системой управления, соответствуют некоторые постоянные значения переменной выхода $y = y^*$ и уровня управляющего воздействия $u = u^*$, обеспечивающие экстремум (максимум или минимум) определенной целевой функции.

Целью управления может быть изменение управляемой переменной y по определенной программе, заданной как функция времени $y_n(t)$, например, программы набора высоты или снижения самолета, программа для станков или обрабатывающих центров в машиностроении, программа изменения температуры в печи и т. п. Для вычисления оптимального управляющего воздействия $u^*(t)$ здесь необходимо знать динамические характеристики объекта управления. При этом не всегда можно найти допустимое управление, обеспечивающее точное следование заданной программе. Существуют методы вычисления оптимального управления $u^*(t)$, когда движение объекта в принятом смысле минимально отличается от заданного.

Более общий случай задания цели управления — это требование оптимизировать некоторый функционал

$$J(y, u) \rightarrow \min_{u \in U} (\max),$$

где U — множество допустимых управлений. В результате решения задачи вариационного исчисления с учетом ограничений на управление и динамики объекта находятся оптимальная траектория $y^*(t)$ и оптимальное управление $u^*(t)$.

Методам оптимизации управлений посвящена отдельная часть книги (см. часть 6).

1.4.2. Компенсация возмущений

Возмущения — воздействия среды на объект, вызывающие отклонения управляемой переменной от заданных значений или программ изменения.

Если о возмущении на объект $f(t)$ имеется полная *априорная* информация, то она может быть учтена при расчете оптимального управления, обеспечивающего желаемое движение объекта.

В некоторых случаях основное возмущение на объект можно измерять непосредственно, т. е. доступна *текущая* информация о причине отклонений. Используя эту информацию, можно компенсировать отклонения управляемой переменной, оказывая на объект дополнительное воздействие. В простейшем случае компенсирующее воздействие суммируется с управлением, как это показано на рис. 1.10. В результате получаем систему, также реализующую принцип разомкнутого управления. Алгоритм обработки текущей информации о возмущении f в компенсаторе K с целью вычисления компенсирующего воздействия δu_f строится на базе информации о характеристиках объекта по каналам управления и возмущения.

Наличие текущей информации о возмущении, т. е. о *причине* отклонений, дало возможность формирования в системе второго (искусственного) пути из точки приложения возмущения к выходу объекта. Наличие такого пути является необходимым условием реализации абсолютной инвариантности управляемой переменной к непосредственно измеряемому возмущению (так называемый *принцип двухканальности*).

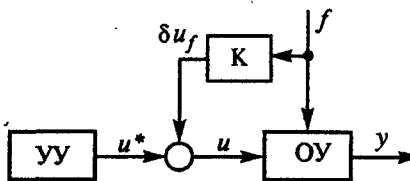


Рис. 1.10. Разомкнутая система с компенсацией возмущения

Для работоспособности системы управления с компенсацией возмущения необходимо, чтобы заданный режим объекта был устойчив, а другие (нескомпенсированные) возмущения были слабыми. Например, судном нельзя управлять длительное время без коррекции курса по выявленному отклонению. Это объясняется тем, что судно является нейтральным в смысле устойчивости объектом, накапливающим отклонения.

1.4.3. Системы управления с обратной связью

Практически невозможно получать и своевременно обрабатывать полную информацию о всех возмущениях, действующих на объект. Для ослабления действия любых возмущений может быть использована текущая информация об отклонениях управляемой переменной, т. е. информация о *следствиях* возмущений. На рис. 1.11 показана схема системы управления, реализующей *принцип обратной связи*. Особенностью системы управления является наличие замкнутого контура причинно-следственных связей.

Достоинство обратной связи заключается в ее универсальности — какие бы возмущения ни действовали на объект, будет выявлено их следствие — отклонение δy управляемой переменной. В регуляторе P на базе этой информации вырабатывается дополнительное управляющее воздействие δu , направленное на уменьшение отклонения δy .

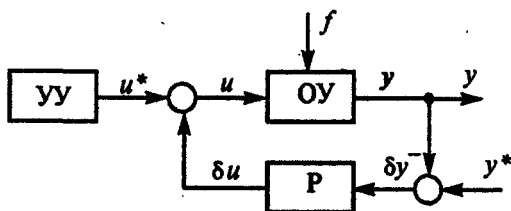


Рис. 1.11. Система с обратной связью

Примером системы с обратной связью является система управления курсом судна (см. рис. 1.2). На рис. 1.3 была приведена структура абстрактной системы управления с обратной связью.

Обратим внимание на то, что для функционирования системы с обратной связью нет необходимости в полной априорной информации о цели управления. Достаточно иметь *текущую* информацию о цели в виде задающего воздействия $y^*(t)$ на входе элемента сравнения. Задание на рассматриваемую систему может генерироваться другими системами, например системами высших уровней иерархии управления. Более того, достаточно иметь текущую информацию только об отклонении $\delta y(t)$ управляемой переменной $y(t)$ от заданных значений $y^*(t)$. Системы, предназначенные для воспроизведения на выходе объекта управления изменяющихся во времени задающих воздействий, называются *следающими*.

Обратная связь является единственным средством стабилизации неустойчивых режимов объектов. Измеряя отклонения от положения

равновесия или в общем случае — от оптимальной траектории, можно оказывать на объект воздействия, направленные на возвращение объекта в исходное состояние.

Наконец, важнейшее свойство отрицательной обратной связи — уменьшение влияния изменений характеристик объекта на свойства системы в целом.

Вышеизложенное позволяет говорить об универсальности принципа обратной связи, которая при правильном применении способна стабилизировать неустойчивые режимы объектов, обеспечивать инвариантность к непосредственно неизмеряемым возмущениям и уменьшать чувствительность к вариациям характеристик объекта и окружающей среды.

Недостатком систем управления с обратной связью является недостижимость абсолютной инвариантности к произвольным воздействиям — для формирования сигнала управления и принципиально необходимы отклонения δu , являющиеся входной информацией управляющего устройства.

1.4.4. Системы с компенсацией параметрических возмущений

Многие объекты изменяют свои свойства либо в зависимости от времени (нестационарность), либо от режима работы. Если доступна априорная информация об изменениях характеристик объекта, то она может быть учтена при расчете оптимального управления. При отсутствии достаточной априорной информации о свойствах объекта управления и/или среды не представляется возможным рассчитать оптимальную программу управления $u^*(t)$ или алгоритм регулятора.

Неопределенность в свойствах объекта или изменение этих свойств можно трактовать как следствие воздействий среды. Если можно непосредственно измерять состояние среды — причины, вызывающие изменение свойств объекта, то текущая информация используется для подстройки параметров управляющих устройств и регуляторов. На рис. 1.12 изображена структура системы с компенсацией параметрических возмущений. Слошными линиями со стрелками обозначены переменные энергетического характера в системе управления с обратной связью, целью которой является воспроизведение на выходе объекта ОУ задающего воздействия u^* с требуемой точностью. Алгоритм регулятора P рассчитан на базе априорной информации о свойствах объекта. Если эти свойства отклоняются от номинальных, качество системы может ухудшаться. На рис. 1.12 наклонная стрелка, «пронизывающая»

объект, условно изображает влияние среды, приводящее к изменению свойств объекта. Если удастся измерять непосредственно воздействия среды F (пунктирная линия на рисунке), то такая текущая информация может быть обработана в блоке компенсации K с целью вычисления новых значений параметров настройки регулятора P , либо для более кардинальных изменений его алгоритма.

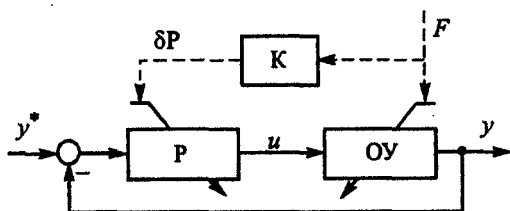


Рис. 1.12. Система с компенсацией параметрических возмущений

Примером систем с компенсацией параметрических возмущений являются некоторые автопилоты. Поскольку известно, как аэродинамические характеристики самолета зависят от высоты полета, то измеряя барометрическое давление, можно подстраивать параметры автопилота.

Отметим, что по отношению к параметрическим воздействиям рассматриваемые системы являются разомкнутыми. Здесь управление осуществляется не по следствию, а по причине, что в принципе позволяет достичь абсолютную инвариантность характеристик таких систем к параметрическим возмущениям.

1.4.5. Адаптивное управление

Несмотря на то что в системах компенсации параметрических возмущений и происходит подстройка параметров, такие системы не принято называть *самонастраивающимися*. Системы управления с самонастройкой или, в общем случае, с адаптацией получают информацию об изменениях свойств объекта путем *текущей* идентификации. Это значит, что каким-либо способом оцениваются следствия параметрических воздействий среды на объект. Текущая информация об отклонениях свойств объекта используется для принятия решения о подстройке параметров управляющих устройств и/или регуляторов. На рис. 1.13 изображена структура адаптивной (самонастраивающейся) системы. Как и в предыдущем случае, следящая система, образованная объектом ОУ и регулятором P , нуждается в подстройке параметров регулятора из-за

изменений свойств объекта. Поскольку причины, вызывающие эти изменения, непосредственно нельзя измерять, остается оценивать их следствие. В устройстве текущей идентификации И путем обработки данных о входе u и выходе y объекта оцениваются отклонения его характеристик. На базе этой информации в блоке адаптации А вычисляются новые значения параметров регулятора Р. Здесь реализуется принцип отрицательной (параметрической) обратной связи — образуется замкнутый контур передачи и обработки информации для ослабления параметрических возмущений среды.

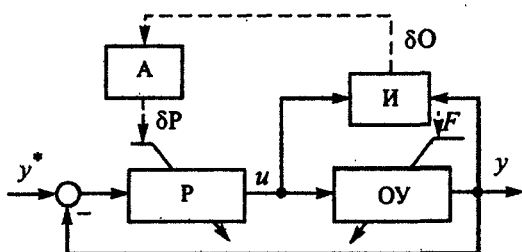


Рис. 1.13. Адаптивная система управления

Заметим, что замкнутый основной контур системы, образованный объектом и регулятором, можно трактовать как объект управления системы более высокого (второго) уровня иерархии. Управляющим воздействием системы второго уровня являются изменения значений параметров регулятора системы первого уровня, а возмущением — параметрические воздействия среды. Роль измерительного элемента в системе второго уровня играет устройство текущей идентификации И, а регулятора — блок настройки А. В блоке А по вариациям характеристик объекта или отклонениям показателей качества вычисляются параметры регулятора Р, из условия сохранения заданных показателей качества системы нижнего уровня иерархии. Это так называемая самонастраивающаяся система управления.

В более общем случае целью адаптивной системы является оптимизация показателей качества, т. е. настройка регулятора связана с поиском экстремума некоторого функционала.

1.5. КЛАССИФИКАЦИЯ СИСТЕМ УПРАВЛЕНИЯ

Предыдущий параграф был, по существу, посвящен классификации систем управления по информационному признаку. Здесь рассмотрим

классификацию по типу сигналов, а также по алгоритмическому и энергетическому признакам.

1.5.1. Классификация систем управления по типу сигнала

Понятия «переменная», «воздействие», «сигнал» в теории управления имеют в основном информационный смысл. Сигналы на входах и на выходах функциональных элементов систем управления рассматриваются как носители информации, а сами элементы — преобразователи сигналов — как устройства передачи и обработки этой информации.

Если информация кодируется уровнями сигналов — функций непрерывного времени (рис. 1.14, *а*), говорят о *непрерывных* системах (системы с непрерывным временем, аналоговые системы). Здесь и время, и уровни всех сигналов принадлежат множеству мощности континуума.

В том случае, когда информация кодируется каким-либо параметром (высота, ширина, фаза) импульсов определенной формы, говорят об *импульсных* системах. Различают импульсные системы с амплитудно-импульсной, широтно-импульсной, фазо-импульсной модуляциями. На рис. 1.14, *б* изображена последовательность прямоугольных импульсов, высота (амплитуда) которых кодирует информацию в моменты времени, кратные периоду T ; это случай амплитудно-импульсной модуляции. В математических моделях импульсных систем используется понятие дискретного времени, принадлежащего счетному множеству, а параметр импульса — амплитуда — принадлежит множеству мощности континуума. Существуют автоматические системы, в которых информация кодируется частотой импульсов.

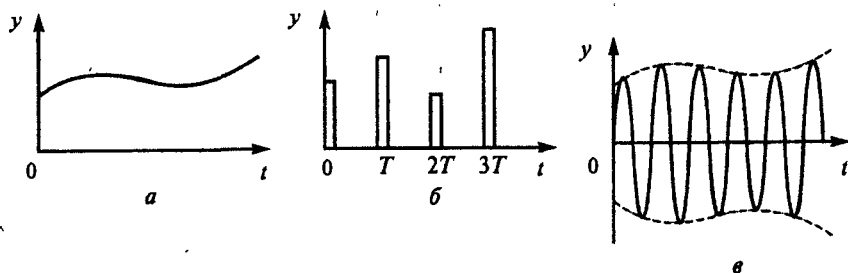


Рис. 1.14. Типы сигналов

В *цифровых* системах управления как и в импульсных происходит квантование времени. Кроме того, в силу конечности разрядной сетки вычислительных устройств имеет место квантование и уровня, т.е. уро-

вень принадлежит конечному множеству (хотя, быть может, и очень большой мощности). Подробнее об импульсных и цифровых системах будет изложено далее в разделах, посвященных теории дискретных систем управления.

Важное место занимают системы управления, в которых информация кодируется амплитудой переменного тока, т.е. огибающей несущего сигнала. Примерами таких систем являются автоматические системы с двигателями переменного тока, системы с радиоканалами с амплитудной модуляцией и др. (рис. 1.14, в).

1.5.2. Классификация систем автоматического регулирования по типу алгоритма

Рассмотрим структуру системы автоматического регулирования с обратной связью (рис. 1.15). Регулятор P на базе информации об ошибке $e(t)$, т.е. отклонении управляемой переменной $y(t)$ от задания $y^*(t)$, формирует по некоторому алгоритму управляющее воздействие $u(t)$ на объект ОУ.

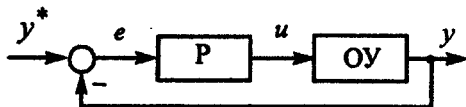


Рис. 1.15. Система автоматического регулирования

Используют различные алгоритмы, или, как еще говорят, законы регулирования. Приведем некоторые из них, полагая, что все переменные являются функциями непрерывного времени.

Пропорциональный закон (П-закон) регулирования имеет ясную логику — чем больше ошибка, тем сильнее воздействие на объект. Если регулятор трактовать как преобразователь сигналов, то математически алгоритм выражается так:

$$u(t) = k_P e(t).$$

Интегральный закон (И-закон) регулирования означает, что уровень воздействия на объект определяется суммарной ошибкой, накопившейся за определенное время:

$$u(t) = k_I \int_0^t e(\tau) d\tau + u_0.$$

Если продифференцировать последнее выражение по времени, получим

$$\frac{du(t)}{dt} = k_I e(t),$$

т.е. ошибка системы определяет скорость изменения управляющего воздействия. Характерной чертой И-закона является то, что управляющее воздействие изменяется до тех пор, пока ошибка не станет равной нулю.

Пропорционально-интегральный закон (ПИ-закон) является комбинацией двух предыдущих:

$$u(t) = k_P e(t) + k_I \int_0^t e(\tau) d\tau.$$

Пропорционально-интегрально-дифференциальный закон (ПИД-закон) формирует управляющие воздействия с учетом и скорости изменения ошибки:

$$u(t) = k_P e(t) + k_I \int_0^t e(\tau) d\tau + k_D \frac{de(t)}{dt}.$$

Этот закон наиболее часто реализуется в системах промышленной автоматики.

1.5.3. Классификация систем управления по энергетическому признаку

Хотя переменные — входы и выходы элементов систем управления — в теории управления рассматриваются как носители информации, не следует забывать, что исполнение принятого в управляющем устройстве решения об оказании управляющего воздействия на объект связано с энергетическими затратами, иногда значительными.

Для оказания управляющего воздействия может использоваться энергия самого управляемого процесса, отбираемая через измерительный элемент. В этом случае говорят о *прямом регулировании*. Широко распространены поплавковые регуляторы уровня прямого действия, например, для карбюраторов автомобильных двигателей (рис. 1.16). Здесь поплавок одновременно осуществляет три операции: измерение уровня, принятие решения об оказании управляющего воздействия и исполнение решения. Аналогичный принцип регулирования уровня воды в котле паровой машины реализовал русский механик И. И. Ползунов.

Английский механик Дж. Уатт (J. Watt) построил прямую систему автоматического регулирования частоты вращения вала паровой машины с центробежным датчиком. Следует заметить, что все приведенные примеры систем прямого регулирования реализуют принцип управления по замкнутому циклу.

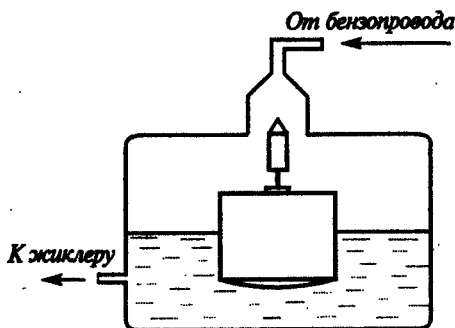


Рис. 1.16. Система регулирования прямого действия

Достоинством систем прямого регулирования является простота и надежность. Они, как правило, органично вписываются в конструкцию управляемых объектов, не требуя посторонних источников энергии. Недостатком этих систем является невысокая точность при необходимости обеспечивать большие усилия для перемещения управляющих органов. Например, при стабилизации уровня жидкости в большом резервуаре с помощью регулятора прямого действия пришлось бы сильно увеличивать объем поплавка. Для точного регулирования частоты вращения паровой машины большой мощности, пар в которую подается через регулирующий клапан с большим усилием перестановки штока, необходимо увеличивать вращающиеся массы центробежного датчика, что приводит к ухудшению качества переходных процессов.

Для повышения точности управления объектами, требующими значительных энергетических затрат на управление, необходимо разделение функций между измерительным и исполнительным элементами систем управления. В системах *непрямого* действия для оказания управляющего воздействия на объект привлекается дополнительный источник энергии. На рис. 1.17 изображена принципиальная схема такой системы для стабилизации уровня жидкости. Измеритель уровня — поплавки — имеет небольшие размеры, так как перемещение подвижного контакта потенциометра Π не требует больших усилий. Напряжение Δu , пропорциональное отклонению уровня от заданного, усиливает-

ется в усилителе У и приводит во вращение двигатель Д. Двигатель через редуктор Р перемещает клапан на трубопроводе притока жидкости. Как усилитель, так и обмотка возбуждения двигателя потребляют электроэнергию дополнительного источника.

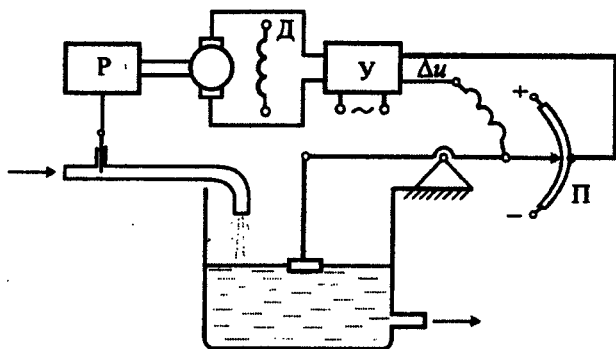


Рис. 1.17. Система регулирования непрямого действия

Достоинством систем непрямого управления является более высокая точность. Недостатком — большая сложность и стоимость, необходимость дополнительных источников энергии, меньшая надежность.

По виду используемой для управления энергии различают электрические, механические, гидравлические, пневматические, электрогидравлические и другие системы автоматического управления.

1.6. ЗАДАЧИ ТЕОРИИ УПРАВЛЕНИЯ

Основными задачами теории управления являются *анализ* и *синтез* систем управления. Под анализом понимается выявление и количественная оценка свойств поведения, а также объяснение свойств систем через характеристики элементов и способ их взаимосвязи. Важнейшими свойствами объектов и систем управления являются: устойчивость движений, инвариантность к возмущениям и робастность (грубость, малая чувствительность). Синтез предполагает выбор элементов и связей между ними таким образом, чтобы система имела требуемое поведение. Таким образом, синтез является задачей, обратной по отношению к анализу.

Ясно, что для систем управления в технике имеют важное значение и такие показатели, как надежность, масса и габариты, энергопотребление, стоимость и эксплуатационные расходы, внешний вид и др. За-

дача обеспечения соответствия этих показателей требованиям решается в процессе проектирования систем управления. Эта задача выходит за рамки теории управления, в которой доминирует поведенческий аспект.

Решение задач анализа и синтеза на реальных объектах возможно в редких случаях. Как правило, это требует много времени, дорого, опасно и не всегда осуществимо. Действительно, управляемый процесс может быть очень длительным (печи и другие тепловые объекты), эксперименты — дорогими (запуск космических ракет) и опасными (ядерные реакторы). Кроме того, объект управления зачастую проектируется вместе с системой управления. Задачи анализа и синтеза в некоторых случаях можно решать с использованием физических моделей (аналогов) объектов.

Теория управления имеет дело с *математическими* (символьными) *моделями*. Построение математических моделей объектов и систем управления также является задачей теории управления и смежных с ней дисциплин. Математические модели позволяют решать задачи анализа и синтеза аналитически (расчетным путем) и путем имитации систем управления на компьютерах.

1.7. МАТЕМАТИЧЕСКИЕ МОДЕЛИ СИСТЕМ УПРАВЛЕНИЯ

Моделирование объектов и систем управления начинается с их выделения из окружающей среды, что, вообще говоря, приводит к искажению изучаемых процессов, так как в природе все явления в той или иной степени взаимосвязаны и взаимообусловлены. Динамические модели объектов и систем управления строятся для объяснения и предсказания их поведения — изменений во времени состояния и наблюдаемых выходных переменных, вызванных внутренними процессами и/или воздействиями среды.

1.7.1. Взаимодействие системы со средой

Рассмотрение причинно-следственного взаимодействия системы управления со средой связано с обособлением собственно системы S и выделением ее связей со средой через переменные входа f и выхода y (рис. 1.18, а). Система оказывается звеном в искусственно разорванной цепи причинно-следственных отношений «среда–система–среда».

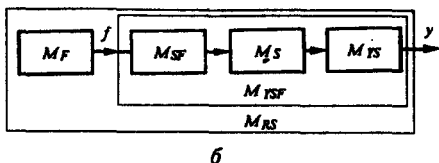
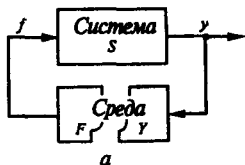


Рис. 1.18. Взаимодействие системы со средой

В теории и расчетной практике объектами исследований оказываются модели собственно систем управления M_S , модели систем со связями со средой M_{YSF} и модели расширенных систем M_{RS} (рис. 1.18, б) [20, 21, 36]. Модели M_S позволяют выявить свойства свободных движений автономных систем, M_{YSF} — свойства каналов передач от входов M_{SF} к выходам M_{YS} при отсутствии информации о переменных входа $f(t)$ — модели среды M_F , а модели M_{RS} привлекаются для изучения вынужденных движений переменных выхода $y(t)$ при адекватных моделях воздействий.

1.7.2. Операторы преобразования переменных

На содержательном уровне объекты и системы управления интерпретируются как устройства получения, передачи и обработки информации. С другой стороны, объекты и системы можно рассматривать как преобразователи сигналов — носителей этой информации. Преобразование сводится к изменению параметров, кодирующих информацию. Свойства системы как преобразователя характеризуются ее *оператором*, отображающим множество функций времени на входе системы на множество функций выхода:

$$y(t) = \mathcal{O}\{f(t)\}.$$

Примерами операторов являются оператор дифференцирования или дифференциальные уравнения

$$\Phi(y, y', y'', \dots, y^{(n)}; u, u', u'', \dots, u^{(m)}) = 0,$$

а также разностные уравнения:

$$\Phi^*(y(k), \Delta y(k), \Delta^2 y(k), \dots, \Delta^n y(k); u(k), \Delta u(k), \dots, \Delta^m u(k)) = 0,$$

где $\Delta y(k) = y(k) - y(k-1)$; $\Delta^2 y(k) = \Delta y(k) - \Delta y(k-1)$ и т.д. — разности первого и высших порядков.

Оператор *линеен*, если обладает свойствами *однородности* и *аддитивности*, т. е.

$$\begin{aligned}O\{af(t)\} &= aO\{f(t)\}; \\O\{f_1(t) + f_2(t)\} &= O\{f_1(t)\} + O\{f_2(t)\}.\end{aligned}$$

Для линейного оператора справедливо:

- при любом усилении (ослаблении) входного воздействия выходная переменная претерпевает точно такое же усиление, не изменяя своей формы;
- реакция на сумму любых входных воздействий равна сумме реакций на эти воздействия.

В общем случае линейной комбинации входных воздействий отвечает та же линейная комбинация соответствующих реакций:

$$O\left\{\sum_i a_i f_i(t)\right\} = \sum_i a_i O\{f_i(t)\}.$$

Свойство линейности оператора, выраженное приведенной формулой, иногда называют *принципом суперпозиции*. Принцип суперпозиции дает возможность выражать реакцию линейной системы на любое воздействие через ее реакцию на определенный вид элементарных воздействий. Для этого произвольное воздействие $f(t)$ представляется как линейная комбинация элементарных воздействий выбранного типа. Зная реакцию линейной системы на элементарные воздействия этого типа, определяется ее реакция на воздействие $f(t)$. Таким образом, линейная система как преобразователь полностью характеризуется ее реакцией на типовое воздействие, называемой временной характеристикой. Хотя в простейших случаях такая форма представления оператора наглядна, ее недостатком является неудобство решения задач анализа и синтеза.

Если не выполняется принцип суперпозиции, то оператор называется *нелинейным*. Разумеется, класс нелинейных операторов много больше класса линейных.

Оператор *стационарен*, если его характеристики инвариантны к времени. Другими словами, при сдвиге во времени входного воздействия без изменения его формы реакция претерпевает такой же сдвиг во времени без изменения своей формы. В ряде случаев модели должны отражать изменение свойств объекта во времени; тогда вводятся для рассмотрения *нестационарные* операторы

$$\mathcal{N}(t) = O\{f(t), t\}.$$

В простейшем случае нестационарность сводится к изменению параметров модели, например, коэффициентов дифференциального уравнения

нения. В общем случае влияние среды приводит к необходимости изменения *структуры* оператора, например, порядка дифференциального уравнения или даже *класса* оператора.

Если вариации оператора происходят много медленнее основных процессов, то вместо нестационарного оператора рассматривают множество стационарных операторов, различающихся значениями параметров. Описание объекта множеством равновероятных операторов содержит *неопределенность*. Если параметры модели заданы с точностью до интервалов значений, то о таких системах говорят, что они *интервальные*.

Оператор может быть *детерминированным* или *стохастичным*. В случае стохастичных операторов параметры представляются как случайные величины и задаются их вероятностными характеристиками.

Объекты управления могут быть с *сосредоточенными* или *распределенными* параметрами. В последнем случае они описываются уравнениями в частных производных (разностях).

1.7.3. Классы моделей

Модель объекта или системы управления принадлежит тому же классу, что и описывающий их оператор преобразования. Разумеется, что можно говорить о классе только математической модели, а не реальной системы.

Таким образом, выделяют следующие признаки классов систем с непрерывным и дискретным временем:

- линейные (Л) или нелинейные (Н);
- стационарные (С) или нестационарные (НС);
- детерминированные (Д) или стохастичные (СД);
- сосредоточенные (конечномерные) (К) или распределенные (бесконечномерные) (БК).

Эти четыре независимых признака бинальтернативны; поэтому можно насчитать всего $2^4 = 16$ классов непрерывных и столько же — дискретных систем.

Простейший класс ЛСДК — линейные стационарные детерминированные конечномерные системы. Они имеют форму обыкновенных линейных дифференциальных (разностных) уравнений с постоянными детерминированными коэффициентами. В математике разработан весьма развитый аппарат анализа этого класса систем. При построении моделей стремятся к их максимальной простоте при достаточной адекватности оригиналам. Поэтому в первом приближении часто ограничиваются описанием объектов управления в классе ЛСДК.

Более сложные классы операторов получаются при введении одного из альтернативных признаков:

ЛСДК; ЛСДК; ЛСДК; ЛСДК.

Для таких систем существует незначительное число общих методов аналитического исследования; в основном, они разработаны только для частных случаев.

Операторы второго уровня сложности получаются введением двух отрицаний

ЛСДК; ЛСДК; ЛСДК; ЛСДК; ЛСДК; ЛСДК.

При трех отрицаниях получаем операторы третьего уровня сложности:

ЛСДК; ЛСДК; ЛСДК; ЛСДК.

Наконец операторы четвертого уровня сложности

ЛСДК

— нелинейные нестационарные стохастические бесконечномерные. Им, например, соответствуют нелинейные дифференциальные уравнения в частных производных с переменными случайными параметрами.

Для систем, описываемых операторами второго и выше уровней сложности, как правило, имеется только единственная возможность их анализа и синтеза путем вычислительных экспериментов.

Если модель системы образована элементами различных классов, то класс системы определяется классом элемента с максимальным числом отрицаний.

1.7.4. Автономные системы

Система называется *автономной*, если на нее не действуют внешние силы, в том числе параметрического типа. Автономные системы, таким образом, стационарны. Изменение их состояния происходит в силу накопленной ранее энергии. На рис. 1.19 модель среды представлена в виде автономной системы, имеющей выход, но не имеющей входов.



Рис. 1.19. Автономная система

Дифференциальные уравнения автономных систем включают переменные системы и их производные, но не содержат переменных, описывающих воздействия среды и имеют постоянные параметры. Это так называемые *однородные дифференциальные уравнения*

$$\Phi(y, y', y'', \dots, y^{(n)}) = 0,$$

дополняемые начальными условиями:

$$y(0), y'(0), \dots, y^{(n-1)}(0).$$

Начальные условия являются следствием предыстории системы и вместе с дифференциальными уравнениями полностью определяют поведение автономной системы. Если система автономна, то ее движения называют *свободными*.

В случае автономных систем с дискретным временем будем иметь однородные разностные уравнения:

$$\Phi^*(y(k), \Delta y(k), \Delta^2(k), \dots, \Delta^n y(k)) = 0.$$

1.7.5. Модели среды

Среда на входе системы моделируется автономными системами — *генераторами* воздействий (см. рис. 1.19) или преобразователями типовых воздействий — *фильтрами*. Распространенными типовыми сигналами, моделирующими детерминированное воздействие, являются единичные импульсная и ступенчатая функции. Примером типового случайного воздействия является так называемый «белый шум». Среда может моделироваться динамической системой того же класса, что и сама система управления. Однако часто рассматриваются детерминированные системы со случайными воздействиями на входе.

1.8. СПОСОБЫ ПОСТРОЕНИЯ МОДЕЛЕЙ

В зависимости от характера и объема априорной информации об объекте исследования выделяют два способа построения моделей систем управления в формах, принятых в теории управления:

- 1 — аналитический;
- 2 — экспериментальный.

1.8.1. Аналитический способ

Аналитический способ применяется для построения моделей объектов хорошо изученной природы. В этом случае имеется вся необходимая информация о свойствах объекта, но она представлена в другой форме.

В результате идеализации физических объектов появляются структурные модели в виде схем с *сосредоточенными компонентами* (рис. 1.20, а). Типичными представителями физических систем, допускающих такое представление, являются электрические и механические объекты. На рис. 1.20, б изображена электрическая схема; рис. 1.20, в представляет собой пример механической поступательной системы.

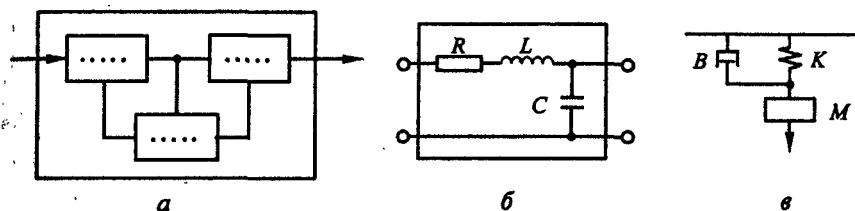


Рис. 1.20. Физические системы с сосредоточенными компонентами (а): электрическая схема (б) и механическая поступательная система (в)

Подобные схемы являются моделями, в которых информация об интересующих свойствах объекта представлена в наглядной форме с использованием графических образов, отражающих физическую природу явлений, устройство и параметры объектов. На таких моделях базируются соответствующие дисциплины, например теоретическая электротехника и теоретическая механика. По существу, принципиальные схемы — стационарные линейные модели с сосредоточенными компонентами (конечномерные), только представлены они с использованием обозначений, принятых соответственно в электротехнике и механике.

Методы теории управления абстрагируются от конкретной природы объектов и оперируют более общими — математическими (символьными) моделями.

Аналитический способ моделирования складывается из этапа построения схемы объекта и ее дальнейшего преобразования в математическое описание требуемой формы. При этом принципиальные проблемы моделирования решаются на первом — неформальном этапе. Вторым этапом, по-существу, оказывается процедура преобразования форм представления моделей. Это позволяет разработать различные компьютерные программы, позволяющие автоматизировать составление уравнений по схемам.

Рассмотрим примеры составления дифференциальных уравнений электрического и механического объектов. Ограничимся классом линейных стационарных моделей.

1.8.2. Электрическая схема

Существует три типа пассивных электрических двухполюсников — резистор R , конденсатор C и катушка индуктивности L — описываемых следующими уравнениями для токов $i(t)$ и напряжений $u(t)$:

$$u_R(t) = Ri_R(t);$$

$$u_C(t) = \frac{1}{C} \int_0^t i_C(\tau) d\tau;$$

$$u_L(t) = L \frac{di_L(t)}{dt}.$$

Активными двухполюсниками электрических схем являются *источник напряжения* и *источник тока*.

Уравнения связи двухполюсников в конкретной схеме выражаются законами Кирхгофа (Kirchhoff G. R.), представляющими собой условия непрерывности токов и равновесия напряжений:

1 — сумма токов в любом узле равна нулю;

2 — сумма напряжений в любом контуре равна нулю.

Рассмотрим пример электрической схемы, изображенной на рис. 1.20, б. Пусть выходом схемы является напряжение на конденсаторе $u_C(t)$. В соответствии с первым законом имеем:

$$i_R = i_L = i_C = i.$$

Второй закон для единственного контура запишется так:

$$u + u_R + u_L + u_C = 0.$$

Выражая напряжения u_R и u_L через u_C :

$$u_R = RC \frac{du_C}{dt}; \quad u_L = LC \frac{d^2 u_C}{dt^2},$$

получим дифференциальное уравнение второго порядка:

$$LC \frac{d^2 u_C}{dt^2} + RC \frac{du_C}{dt} + u_C = -u.$$

1.8.3. Механическая поступательная система

Пассивными двухполюсниками механических схем являются *механическое сопротивление* B , *масса* M и *упругость* K , описываемые следующими уравнениями для сил f и перемещений x или скоростей v :

$$f_B(t) = B \frac{dx_B(t)}{dt} = Bv_B(t);$$

$$f_M(t) = M \frac{d^2 x_M(t)}{dt^2} = M \frac{dv_M(t)}{dt};$$

$$f_K(t) = Kx_K(t) = K \int_0^t v_K(\tau) d\tau.$$

Идеальными источниками механической энергии являются *источник скорости* и *источник силы*.

Уравнения связей механических двухполюсников выражают условия равновесия сил и непрерывности перемещений (скоростей).

Рассмотрим пример механической поступательной системы, изображенной на рис. 1.20, в. По этому рисунку строят схему системы далее в соответствии с приведенными выше уравнениями механических двухполюсников и уравнений связей записывают дифференциальное уравнение для перемещений:

$$M \frac{d^2 x}{dt^2} + B \frac{dx}{dt} + Kx = 0.$$

В этом однородном уравнении отсутствует правая часть, описывающая внешнее воздействие на механическую систему, т.е. она автономна. Свободные движения автономной системы являются следствием нулевых начальных условий, например начального смещения $x(0)$ от равновесного состояния.

Обратим внимание на то, что на рис. 1.20 связи между компонентами не являются направленными, а сами двухполюсники не имеют входов и выходов. Между переменными, характеризующими состояние схем, нет объективных причинно-следственных отношений. Можно считать, что напряжение на резисторе является причиной протекающего по нему тока, но можно сказать и иначе — протекающий ток является причиной падения напряжения на резисторе. Вместе с тем, взаимодействие между схемой и средой, моделируемой источниками, имеет причинно-следственный, направленный характер.

При моделировании объектов различной природы — электрической, механической поступательной и вращательной, гидравлической или пневматической и др., а также смешанной природы, например, электромеханической (двигатели, генераторы) — могут быть выделены аналогичные пассивные и активные компоненты. Дальнейшей абстракцией при построении моделей физических объектов с сосредоточенными компонентами является *полюсный граф* [70]. Эти универсальные топологии

логические модели позволяют унифицировать составление уравнений. Специфика предметной области проявляется только на этапе построения схемы и полюсного графа, а также на заключительном этапе интерпретации результатов анализа и синтеза.

В заключение заметим, что в результате аналитического моделирования сложных объектов получают системы уравнений в *непричинно-следственной форме* относительно внутренних переменных [8].

При проектировании систем управления, когда некоторые элементы не существуют в натуре, аналитический метод построения моделей оказывается единственно возможным.

1.8.4. Экспериментальный способ

Если свойства объекта познаны в недостаточной степени, либо происходящие явления слишком сложны для аналитического описания, для построения математических моделей реально существующих объектов применяется *экспериментальный способ*. Этот способ заключается в активных экспериментах над объектом или в пассивной регистрации его поведения в режиме нормальной эксплуатации (рис. 1.21, а). В результате обработки данных наблюдений получают модели в требуемой форме. Совокупность этих операций объединяется термином «*идентификация объекта*». В результате идентификации получают модели *вход-выход* (рис. 1.21, б).

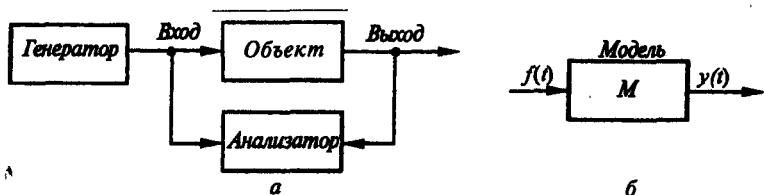


Рис. 1.21. Экспериментальное исследование системы (а) и модель вход-выход (б)

Очевидно, модель зависит не только от свойств объекта, но также от входных сигналов, их разнообразия.

Практически об идентифицируемом объекте всегда имеется какая-то априорная информация, т. е. он не является «черным ящиком». Это дает возможность комбинировать оба способа — вначале аналитически строить структуру модели и определять начальные приближенные значения параметров, а далее обработкой экспериментальных данных уточнять их значения.

1.9. ОСОБЕННОСТИ СТРУКТУРНЫХ МОДЕЛЕЙ СИСТЕМ УПРАВЛЕНИЯ

Особенностью математических моделей систем управления является то, что они не только содержат априорную информацию о ее динамических свойствах, необходимую для изучения поведения системы в целом, но также отражают процессы получения и обработки текущей информации о цели системы, состоянии объекта и воздействиях среды для принятия решения по оказанию на объект надлежащего управляющего воздействия. При построении моделей систем управления и выборе форм их представления учитываются не только динамические, но и информационные, а также алгоритмические аспекты проблемы. Поскольку модели элементов и систем являются основным материалом в задачах анализа и синтеза (исходными данными и результатами), то им и алгоритмам их преобразования в теории управления отводят важное место.

1.9.1. Модели систем управления с раскрытой причинно-следственной структурой

Понятие модели системы управления неотделимо от понятия *структуры*. Под структурой систем управления понимают *причинно-следственные взаимосвязи элементов (подсистем) направленного действия*. Именно ориентированность элементов и их взаимосвязей отличает модели систем управления от структурных моделей физических систем вообще.

При построении моделей с раскрытой причинно-следственной структурой (рис. 1.22) объект или систему предварительно расчленяют на элементы направленного действия и рассматривают их как преобразователи сигналов. Элементы, как правило, выделяются по функциональному признаку, причем сами эти функции понимаются в контексте операций управления: объект управления; измерительные, преобразовательные и усилительные элементы; управляющее устройство; исполнительный механизм; управляющий орган. Примером такого расчленения является приведенная ранее функциональная схема (см. рис. 1.3). Далее для каждой части строится своя модель, а затем модели частей связывают между собой таким же образом, как соединялись сами части.

Если части системы образуют контуры, то моделирование по частям встречается с принципиальной проблемой: не зная свойств частей, нельзя описать сигналы на их входах; не зная сигналов, нельзя пра-

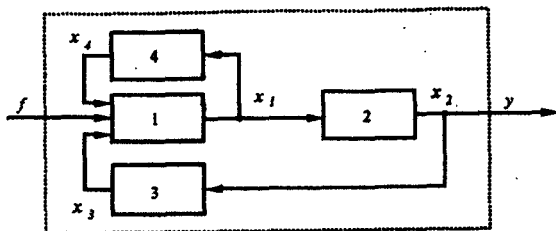


Рис. 1.22. Система управления с раскрытой структурой

вильно идентифицировать отдельные части. Кроме того, возникают известные трудности и при принятии допущения об однонаправленности частей.

Достоинство моделирования по частям — модели содержат в общем случае большую информацию о системе; они вскрывают механизм преобразования входов в выходы.

1.9.2. Иерархические модели систем управления

С точки зрения специалиста по управлению, модели вход-выход (рис. 1.21, б) и структурные модели физических систем (рис. 1.20), имеют *нулевой уровень* причинно-следственной интеграции. Модель, изображенная на рис. 1.22, образована ориентированной взаимосвязью подсистем нулевого уровня и, следовательно, имеет *первый уровень*. Дальнейшее раскрытие структур подсистем приводит к *многоуровневым* (иерархическим) моделям (рис. 1.23).

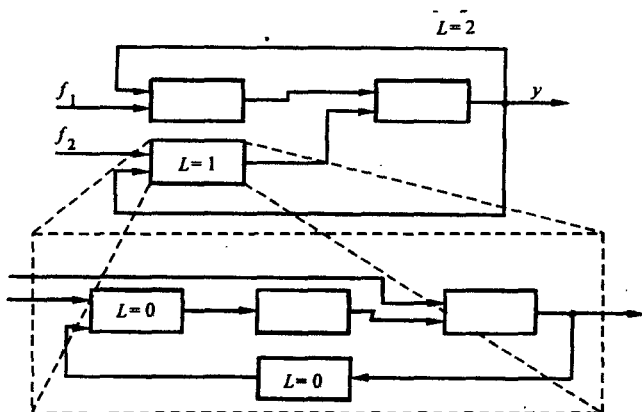


Рис. 1.23. Модель второго уровня

Раскрытие структуры системы или подсистемы означает переход к нижележащему уровню исследования. В зависимости от глубины раскрытия структуры получаются модели различных уровней причинно-следственной интеграции $L = 0, 1, 2, \dots$ [20, 21, 36]. На рис. 1.23 иллюстрируется модель второго уровня ($L = 2$), построенная на подсистемах первого уровня.

Иерархический подход к моделированию позволяет разработать методы исследования и проектирования сложных систем управления. В силу одного из принципов системного подхода — *принципа рекуррентного объяснения* [85] — поведение системы L -го уровня объясняется свойствами подсистем непосредственно нижележащего — $(L-1)$ -го уровня и особенностями их взаимосвязей. Поэтому при разработке методов анализа и синтеза можно ограничиться рассмотрением моделей нулевого и первого уровней.

ЧАСТЬ 2

ТЕОРИЯ ЛИНЕЙНЫХ НЕПРЕРЫВНЫХ СИСТЕМ УПРАВЛЕНИЯ

Глава 2

ЛИНЕЙНЫЕ МОДЕЛИ И ХАРАКТЕРИСТИКИ СИСТЕМ УПРАВЛЕНИЯ

Рассмотрим различные формы представления конечномерных линейных непрерывных стационарных детерминированных моделей систем управления.

2.1. МОДЕЛИ ВХОД-ВЫХОД

Основными формами представления операторов преобразования входных переменных $f(t)$ в переменные выхода $y(t)$ являются: *дифференциальные уравнения, передаточные функции, временные и частотные характеристики*. Для одномерных систем переменные $f(t)$ и $y(t)$ являются скалярами. Эти и некоторые другие представления операторов рассматриваемого класса моделей могут быть приняты за основу задания динамических свойств в терминах вход-выход. Если для конкретных исследований та или иная форма оказывается более предпочтительной, ставится и решается задача перехода от одной формы к другой, например, построения временных и частотных характеристик по дифференциальному уравнению или передаточной функции.

2.1.1. Дифференциальные уравнения

Обыкновенное линейное дифференциальное уравнение n -го порядка с постоянными коэффициентами обычно записывается так:

$$a_n \frac{d^n y}{dt^n} + \dots + a_1 \frac{dy}{dt} + a_0 y = b_m \frac{d^m f}{dt^m} + \dots + b_1 \frac{df}{dt} + b_0 f. \quad (2.1)$$

Если ввести оператор дифференцирования по времени

$$p \equiv d / dt,$$

то уравнение (2.1) запишется в компактном виде

$$A(p)y(t) = B(p)f(t), \quad (2.2)$$

где $A(p) = a_n p^n + \dots + a_1 p + a_0$; $B(p) = b_m p^m + \dots + b_1 p + b_0$ — операторные полиномы. Дифференциальное уравнение дополняется начальными условиями $y(0), y'(0), \dots, y^{(n-1)}(0)$.

2.1.2. Передаточные функции

Передаточная функция равна отношению изображений по Лапласу переменных выхода и входа при нулевых начальных условиях

$$W(s) = \frac{Y(s)}{F(s)},$$

где интегральное преобразование Лапласа (P. Laplace) определяется так:

$$Y(s) = L\{y(t)\} = \int_0^{\infty} y(t) e^{-st} dt;$$

$$F(s) = L\{f(t)\} = \int_0^{\infty} f(t) e^{-st} dt.$$

Преобразуя дифференциальное уравнение (2.1) при нулевых начальных условиях, получаем алгебраическое уравнение для изображений

$$A(s)Y(s) = B(s)F(s).$$

Отсюда следует, что передаточная функция легко записывается по дифференциальному уравнению

$$W(s) = \frac{B(s)}{A(s)} \quad (2.3)$$

наоборот, по передаточной функции сразу записывается дифференциальное уравнение.

Зная передаточную функцию и изображение переменной входа, легко найти изображение выхода

$$Y(s) = W(s)F(s).$$

Пример 1. Пусть система описывается дифференциальным уравнением второго порядка

$$a_2 \frac{d^2 y}{dt^2} + a_1 \frac{dy}{dt} + a_0 y = b_0 f; \quad y(0), y'(0).$$

Преобразуем уравнение по Лапласу, для чего воспользуемся свойством линейности оператора преобразования L , а также теоремой о дифференцировании оригинала

$$a_2(s^2 Y(s) - sy(0) - y'(0)) + a_1(sY(s) - y(0)) + a_0 Y(s) = b_0 F(s).$$

Последнее уравнение перепишем в другом виде:

$$(a_2 s^2 + a_1 s + a_0) Y(s) = b_0 F(s) + a_2 s y(0) + a_2 y'(0) + a_1 y(0).$$

При нулевых начальных условиях $y(0) = y'(0) = 0$ отношение изображений, т. е. передаточная функция равна

$$W(s) = \frac{b_0}{a_2 s^2 + a_1 s + a_0}.$$

Оператор, связывающий вход и выход, можно задать коэффициентом и множествами нулей (корней полинома числителя) $z_j; j = 1, \dots, m$, и полюсов (корней полинома знаменателя) $p_i; i = 1, \dots, n$, передаточной функции (2.3)

$$W(s) = \frac{b_m \prod (s - z_j)}{a_n \prod (s - p_i)}. \quad (2.4)$$

В отличие от *полиномиальной* формы (2.3), эту форму задания передаточных функций иногда называют *факторизованной*.

Вводится понятие *структуры оператора* преобразования. Для дифференциального уравнения n -го порядка (2.1) и передаточной функции (2.3) задание структуры означает задание целых чисел — степеней $n = \deg A$ и $m = \deg B$ полиномов A и B .

Параметрами оператора являются коэффициенты полиномов.

2.1.3. Временные характеристики

Временные характеристики являются одной из форм представления операторов преобразования переменной $f(t)$ в переменную $y(t)$. *Импульсная переходная функция* или *функция веса* $w(t)$ — реакция системы на единичный идеальный импульс $\delta(t)$ при нулевых начальных условиях. Переменная выхода определяется как интеграл свертки

$$y(t) = \int_0^t w(\tau) f(t - \tau) d\tau, \quad (2.5)$$

т. е. в этом случае оператор преобразования имеет форму интегрального уравнения. Другая, часто употребляемая временная характеристика — *переходная характеристика* $h(t)$ — реакция системы на единичную ступенчатую функцию $1(t)$ при нулевых начальных условиях. На рис. 2.1 приведен примерный вид временных характеристик для системы второго порядка.

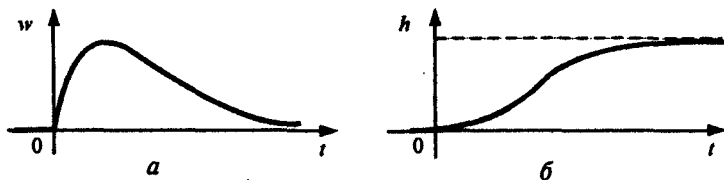


Рис. 2.1. Временные характеристики:
 а — функция веса; б — переходная характеристика

2.1.4. Частотные характеристики

Частотные характеристики элементов и систем представляют собой зависимость параметров установившихся реакций на гармонические сигналы всех частот и единичных амплитуд. В линейных системах форма и частота установившейся реакции совпадают с формой и частотой сигнала на входе. Комплексная частотная характеристика $W(j\omega)$ дает возможность определить амплитуду $R(\omega)$ и фазу $\varphi(\omega)$ гармонического сигнала на выходе системы по значению частоты:

$$W(j\omega) = R(\omega)e^{j\varphi(\omega)} = P(\omega) + jQ(\omega), \quad (2.6)$$

где $R(\omega) = \text{mod} W(j\omega)$ и $\varphi(\omega) = \arg W(j\omega)$ — амплитудная и фазовая частотные характеристики, а $P(\omega) = \text{Re} W(j\omega)$ и $Q(\omega) = \text{Im} W(j\omega)$ — вещественная и мнимая частотные характеристики. На рис. 2.2 изображен пример годографа $W(j\omega)$, называемого *амплитудно-фазовой характеристикой* (АФХ). Реальные объекты с повышением частоты хуже пропускают сигналы — ослабляют амплитуду и вносят отрицательный фазовый сдвиг.

Амплитудно-частотные характеристики удобно представлять в логарифмическом масштабе:

$$L(\omega) = 20 \lg R(\omega).$$

Если частота изменяется в логарифмическом масштабе, то логарифмические амплитудно-частотные характеристики (ЛАЧХ) во многих практически важных случаях мало отличаются от прямолинейных асимптот с наклонами, кратными 20 дБ/дек. На рис. 2.3 приведен примерный вид асимптотической ЛАЧХ; пунктирная линия — точная ЛАЧХ. Там же указаны наклоны асимптот в дБ/дек.

Передаточные функции, все нули и полюсы которых находятся в левой полуплоскости, называют *минимально-фазовыми*. Такие передаточные функции соответствуют меньшим по модулю фазовым сдвигам по сравнению с любыми другими передаточными функциями, имею-

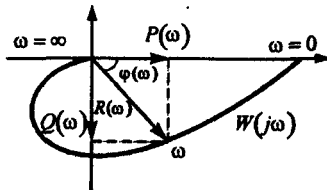


Рис. 2.2. Пример АФХ

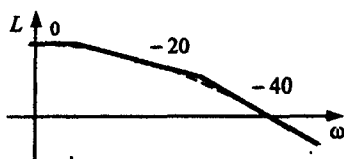


Рис. 2.3. Пример ЛАЧХ

щими ту же ЛАЧХ, но часть нулей и/или полюсов справа от мнимой оси. Для примера на рис. 2.4 изображены АФХ, ЛАЧХ и ЛФЧХ для двух передаточных функций:

$$W_1(s) = Ts + 1; \quad W_2(s) = Ts - 1.$$

Вторая из них имеет правый нуль; ее ЛФЧХ φ_2 имеет большие значения, чем ЛФЧХ φ_1 минимально-фазовой передаточной функции $W_1(s)$.

Для минимально-фазовых передаточных функций справедлива формула Боде (Н. W. Bode) [16], позволяющая вычислять значения фазовой характеристики $\varphi(\theta)$ по крутизне ЛАЧХ $L(\theta)$:

$$\varphi(\theta) = \frac{\pi}{40} \frac{dL(\theta)}{d\theta} + \delta\varphi,$$

где $\theta = \lg \omega$; $\delta\varphi$ — поправочная функция, зависящая от $L(\theta)$ на всем диапазоне частот. Если в окрестности рассматриваемого значения θ ЛАЧХ $L(\theta)$ мало отличается от асимптоты, т.е. нет резких изменений наклона, то $\delta\varphi \approx 0$. Например, если наклон ЛАЧХ близок к 20 дБ/дек, то значения фазовых сдвигов близки к $\pi/2$ рад.

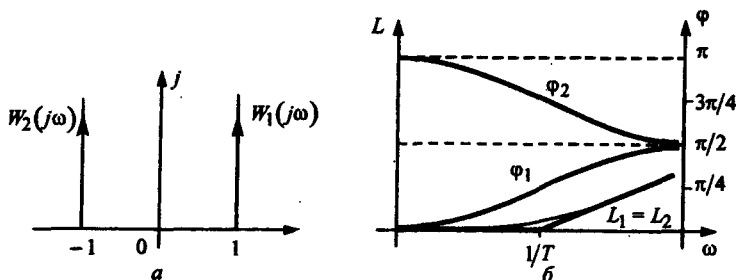


Рис. 2.4. Примеры АФХ (а), ЛАЧХ и ЛФЧХ (б) минимально-фазовой и неминимально-фазовой передаточных функций

2.1.5. Преобразование форм представления моделей вход-выход

Несмотря на то что любая из форм представления операторов может быть принята за основу задания динамических свойств систем, для конкретных исследований та или иная форма оказывается более рациональной. Возникает необходимость перехода от одной формы к другой. Многие задачи анализа связаны с преобразованием формы представления оператора. В ряде случаев эта процедура составляет наиболее трудоемкий этап анализа — построение частной модели — приведение к форме, позволяющей непосредственно вычислить показатели качества и вывести суждение о соответствии поведения системы заданным требованиям (например, построение временных или частотных характеристик системы управления).

Переходы между различными формами представления операторов удобно рассматривать как дуги орграфа, вершинам которого соответствуют формы представления, как это изображено на рис. 2.5.

Наиболее прост формальный переход путем замены оператора дифференцирования $p \equiv d/dt$ на комплексный аргумент s от дифференциального уравнения (2.2) к передаточной функции (2.3) и обратно. Осуществляя переход к передаточным функциям, следует избегать сокращения общих делителей полиномов числителей и знаменателей, т.е. диполей рациональных функций. Такое сокращение приводит к потере части собственных составляющих движения при ненулевых предначальных условиях (составляющих свободных движений).

Пунктирные линии графа взаимосвязи (см. рис. 2.5) отвечают переходам, рассматриваемым обычно в задачах идентификации. По временным и/или частотным характеристикам, полученным экспериментально, оценивают параметры передаточных функций или ординаты характеристик иного типа. Такие переходы оказываются неоднозначными, а их результаты зависят от выбора структуры оператора и алгоритма обработки данных.

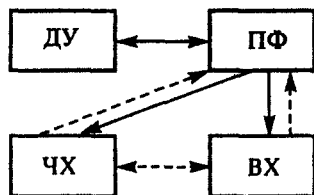


Рис. 2.5. Орграф взаимосвязи форм представления операторов
ДУ — дифференциальное уравнение;
ЧХ — частотные характеристики;
ПФ — передаточная функция;
ВХ — временные характеристики

2.2. ПОСТРОЕНИЕ ВРЕМЕННЫХ ХАРАКТЕРИСТИК

Временные характеристики — импульсная переходная функция $w(t)$ и переходная характеристика $h(t)$ — могут быть получены экспериментально, если удастся подать на вход объекта воздействие в виде доста-

точно узкого импульса с необходимой амплитудой или ступенчатой функции времени. Последнее более реально; функцию веса $w(t)$ впоследствии можно получать дифференцированием функции $h(t)$.

Статистические методы непараметрической идентификации позволяют оценить ординаты функции веса $w(t)$ путем обработки данных вход-выход объекта в виде случайных сигналов, быть может, имеющих место в режиме нормальной эксплуатации (корреляционный анализ).

Существуют методы построения временных характеристик по частотным, базирующиеся на обратном преобразовании Фурье (I. Fourier).

В случае, когда исходная информация об объекте представлена в форме дифференциального уравнения (2.1), временные характеристики получают его решением.

В классической теории автоматического управления для решения дифференциальных уравнений часто привлекают так называемый операторный метод, связанный с преобразованием Лапласа. Метод особенно удобен в случае типовых воздействий в виде обобщенных функций и позволяет легко учесть ненулевые начальные условия.

Пусть дано дифференциальное уравнение n -го порядка звена или системы управления (2.2). Необходимо получить выражения для импульсной переходной функции (функции веса) $w(t)$, переходной характеристики $h(t)$, а также для реакции в случае воздействия общего вида. Пусть изображение воздействия по Лапласу на входе системы или звена представляет собой дробно-рациональную функцию от s

$$F(s) = \frac{B_F(s)}{A_F(s)}.$$

Если преобразовать по Лапласу дифференциальное уравнение n -го порядка при ненулевых предначальных условиях, то после разрешения полученного алгебраического уравнения относительно изображения переменной выхода имеем

$$Y(s) = \frac{B(s)}{A(s)} \frac{B_F(s)}{A_F(s)} + \frac{A_n(s)}{A(s)}. \quad (2.7)$$

Здесь полином $A_n(s)$ определяется предначальными условиями. Если все предначальные условия нулевые, то изображение выхода равно:

$$Y(s) = W(s)F(s) = \frac{B(s)}{A(s)} \frac{B_F(s)}{A_F(s)},$$

где $W(s)$ — передаточная функция.

Искомое решение — переменная на выходе системы (оригинал) — получается путем обратного преобразования Лапласа

$$y(t) = L^{-1}\{Y(s)\} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} Y(s)e^{st} ds, \quad (2.8)$$

где c — абсцисса сходимости. Формула обращения Римана–Меллина (B. Riemann — H. Mellin) устанавливает однозначное соответствие между оригиналом и изображением в точках непрерывности оригинала. Имеются алгоритмы и программы, позволяющие вычислять интеграл (2.8) при произвольных функциях $Y(s)$. Практическое вычисление оригинала $y(t)$ удобно производить, основываясь на теореме о вычетах, согласно которой значение интеграла (2.8) может быть представлено суммой вычетов подынтегральной функции

$$y(t) = \sum_{i=1}^{n_Y} \text{Res} Y(s) e^{s_i t} \Big|_{s=s_i},$$

где $\text{Res } Y(s)$ — вычет функции $Y(s)$ в полюсе s_i ; $i = 1, \dots, n_Y$; n_Y — число простых полюсов изображения $Y(s)$. При $t < 0$ функция $y(t) = 0$.

Для обыкновенных линейных дифференциальных уравнений и типовых воздействий изображение $Y(s)$ является дробно-рациональной функцией, которую можно представить в виде суммы простейших дробей

$$Y(s) = \frac{B_Y(s)}{A_Y(s)} = \sum_{i=1}^{n_Y} \frac{B_Y(s_i)}{A'_Y(s_i)} \frac{1}{s - s_i} = \sum_{i=1}^{n_Y} \frac{C_i}{s - s_i}, \quad (2.9)$$

где A'_Y — производная полинома A_Y по s ; s_i — простые полюсы;

$$C_i = \text{Res} Y(s) = \frac{B_Y(s_i)}{A'_Y(s_i)}.$$

Оригинал $y(t)$ в соответствии с разложением (2.9) имеет вид

$$y(t) = \sum_{i=1}^{n_Y} C_i e^{s_i t}.$$

Импульсная переходная функция (функция веса) $w(t)$ представляет собой реакцию системы на δ -функцию при нулевых начальных условиях. Поскольку изображение δ -функции

$$L\{\delta(t)\} = 1,$$

функция веса представляет собой обращение по Лапласу передаточной функции

$$w(t) = L^{-1}\{W(s)\}.$$

Разложение передаточной функции на сумму простейших дробей в случае простых полюсов s_i ; $i = 1, \dots, n$, имеет вид

$$W(s) = \sum_{i=1}^n \frac{C_i}{s - s_i}, \quad (2.10)$$

где C_i — коэффициенты разложения (вычеты):

$$C_i = \frac{B(s_i)}{A'(s_i)}. \quad (2.11)$$

Пример 2. Найдем с помощью формул (2.10) и (2.11) выражение для функции веса по передаточной функции

$$W(s) = \frac{2}{s^2 + 3s + 2}. \quad (2.12)$$

Полосы передаточной функции $s_1 = -1$; $s_2 = -2$. Разложение (2.12) на сумму простейших дробей имеет вид

$$W(s) = \frac{2}{s+1} - \frac{2}{s+2}.$$

Обратное преобразование Лапласа дает

$$w(t) = 2e^{-t} - 2e^{-2t}.$$

Переходная характеристика $h(t)$ представляет собой реакцию системы на единичную ступенчатую функцию $1(t)$ при нулевых начальных условиях.

Поскольку

$$L\{1(t)\} = 1/s,$$

имеем

$$h(t) = L^{-1}\{H(s)\} = L^{-1}\{W(s)/s\}.$$

Полосами изображения реакции являются полюс воздействия $s_1 = 0$ и полюсы передаточной функции. Легко убедиться, что

$$w(t) = dh(t)/dt,$$

$$h(t) = \int_0^t w(\tau) d\tau.$$

Пример 3. Получим выражение для переходной характеристики системы с передаточной функцией (2.12). Разложение изображения $H(s)$ на сумму простейших дробей дает

$$H(s) = \frac{B_H(s)}{A_H(s)} = \frac{2}{s(s^2 + 3s + 2)} = \frac{C_1}{s} + \frac{C_2}{s+1} + \frac{C_3}{s+2},$$

где

$$C_1 = \frac{B_H(0)}{A'_H(0)} = \frac{2}{3s^2 + 6s + 2} \Big|_{s=0} = 1;$$

$$C_2 = \frac{2}{3s^2 + 6s + 2} \Big|_{s=-1} = -2;$$

$$C_3 = \frac{2}{3s^2 + 6s + 2} \Big|_{s=-2} = 1.$$

Следовательно, переходная характеристика описывается функцией:

$$h(t) = 1 - 2e^{-t} + e^{-2t}.$$

В общем случае произвольного воздействия разложение изображения переменной выхода (2.7) запишется как

$$y(t) = \sum_{i=1}^n \frac{B(s_i)B_F(s_i)}{A'(s_i)A_F(s_i)} e^{s_i t} + \sum_{k=1}^{n_F} \frac{B(s_k)B_F(s_k)}{A(s_k)A_F(s_k)} e^{s_k t} + \sum_{i=1}^n \frac{A_n(s_i)}{A'(s_i)} e^{s_i t}, \quad (2.13)$$

где $s_i, i = 1, \dots, n$ — полюсы передаточной функции $W(s)$; $s_k, k = 1, \dots, n_F$ — полюсы изображения воздействия $F(s)$. Принято, что $s_k \neq s_i$, т.е. полюсы воздействия не равны полюсам передаточной функции (нет обобщенного резонанса).

В выражении (2.13) первая группа слагаемых определяет переходную составляющую вынужденного движения $y_{\text{пер}}(t)$; вторая группа — установившуюся составляющую вынужденного движения $y_{\text{уст}}(t)$, а третья — свободные движения $y_{\text{св}}(t)$:

$$y(t) = y_{\text{пер}}(t) + y_{\text{уст}}(t) + y_{\text{св}}(t).$$

Установившееся вынужденное движение $y_{\text{уст}}(t)$ обусловлено полюсами изображения воздействия s_k ; переходная составляющая вынужденного движения $y_{\text{пер}}(t)$ образуется из-за ненулевых посленаачальных условий (изменение начальных условий приложением в момент времени $t = 0$ конкретного воздействия) и определяется полюсами передаточной функции; свободные движения $y_{\text{св}}(t)$ имеют место при ненулевых предначальных условиях и также определяются полюсами передаточной функции.

Если анализируется автономная система управления M_s , представленная в форме однородного дифференциального уравнения

$$A(p)y(t) = 0; \quad y(0),$$

то его решение имеет вид

$$y(t) = y_{\text{св}}(t) = \sum_{i=1}^n \frac{A_n(s_i)}{A'(s_i)} e^{s_i t}. \quad (2.14)$$

Если изображение $Y(s)$ имеет кратные полюсы, то вместо формул (2.13), (2.14) записываются более сложные выражения [28].

2.3. ПОСТРОЕНИЕ ЧАСТОТНЫХ ХАРАКТЕРИСТИК

Частотные характеристики (2.6) — амплитудную $R(\omega)$ и фазовую $\varphi(\omega)$ — можно получать экспериментальным путем, если удастся подавать на вход устойчивого объекта гармонические воздействия различных частот из диапазона, существенного для выявления требуемых свойств объекта. Статистические методы непараметрической идентификации (спектральный анализ) позволяют оценить значения частотных характеристик путем обработки временных последовательностей на входе и выходе объекта.

Частотные характеристики можно получить по временным характеристикам с помощью преобразования Фурье.

В том случае, когда исходная информация об объекте представлена в форме дифференциального уравнения (2.1), частотные характеристики строят расчетным путем.

Рассмотрим переходы от дифференциального уравнения n -го порядка (2.1) и передаточной функции (2.3) к частотным характеристикам (см. рис. 2.2 и 2.3).

Установившиеся реакции линейной системы на гармоническое воздействие единичной амплитуды

$$f(t) = \exp\{j\omega t\} = \cos \omega t + j \sin \omega t$$

соответствуют частному решению неоднородного дифференциального уравнения (2.2). Будем искать частное решение в виде

$$y(t) = R(\omega) \exp\{j(\omega t + \varphi(\omega))\},$$

где $R(\omega)$, $\varphi(\omega)$ — амплитуда и фаза, в общем случае зависящие от частоты.

Учтем, что

$$p f(t) = \frac{d}{dt} e^{j\omega t} = (j\omega) e^{j\omega t},$$

$$p^2 f(t) = \frac{d^2}{dt^2} e^{j\omega t} = (j\omega)^2 e^{j\omega t}, \dots,$$

а также

$$p y(t) = R(\omega) e^{j\varphi(\omega)} (j\omega) e^{j\omega t},$$

$$p^2 y(t) = R(\omega) e^{j\varphi(\omega)} (j\omega)^2 e^{j\omega t}, \dots$$

Подставим эти соотношения в неоднородное дифференциальное уравнение (2.2), записанное в операторной форме

$$A(j\omega)R(\omega)e^{j\varphi(\omega)}e^{j\omega t} = B(j\omega)e^{j\omega t}.$$

После деления обеих частей на $\exp\{j\omega t\}$ можно записать

$$R(\omega)e^{j\varphi(\omega)} = \frac{B(j\omega)}{A(j\omega)} = W(j\omega).$$

Таким образом, амплитудно-частотная характеристика находится как модуль

$$R(\omega) = \text{mod } W(j\omega),$$

а фазо-частотная характеристика — как аргумент

$$\varphi(\omega) = \arg W(j\omega)$$

комплексной функции действительного аргумента $W(j\omega)$.

Одновременно получаем переход от передаточной функции к частотным характеристикам. Комплексная частотная характеристика получается заменой $s = j\omega$ аргумента передаточной функции:

$$W(s)|_{s=j\omega} = W(j\omega) = R(\omega)e^{j\varphi(\omega)}; \quad \omega \geq 0.$$

В общем случае s может принимать значения на любом контуре комплексной плоскости.

Вычисление значений частотных характеристик для конкретного $s = j\omega$ (а в общем случае $s = \alpha + j\omega$) сводится к вычислению значений полиномов $B(s)$ и $A(s)$ с последующим делением полученных комплексных чисел. При этом получаются значения вещественной $P(\omega)$ и мнимой $Q(\omega)$ частотных характеристик. Значения амплитудно-частотной характеристики вычисляются по формуле

$$R(\omega) = \sqrt{P^2(\omega) + Q^2(\omega)}.$$

Трудности возникают при расчете значений фазо-частотной характеристики по формуле

$$\varphi(\omega) = \text{Arctg} \frac{Q(\omega)}{P(\omega)} + k\pi; \quad k = 0, \pm 1, \pm 2, \dots \quad (2.15)$$

Значения $\varphi(\omega)$ получаются на интервале $(-\pi, \pi)$, поэтому в случае систем высокого порядка для определения истинных значений фазовых сдвигов принимается предположение о том, что в пределах выбранного шага частот $\varphi(\omega)$ не изменяется на $\pm\pi$, т.е. корни полиномов $B(s)$ и $A(s)$ располагаются достаточно далеко от мнимой оси.

Вообще говоря, соотношение (2.15) не определяет аргумента $\varphi(\omega)$ комплексного числа $W(j\omega)$, так как ему вместе с φ удовлетворяет и $(\varphi + \pi)$. Однако из-за непрерывности фазо-частотной характеристики $\varphi(\omega)$ [61], принимающей отличные от нуля значения, она однозначно характеризуется текущим

$$\operatorname{tg} \varphi(\omega) = Q(\omega) / P(\omega); \quad \omega_{\min} < \omega < \omega_{\max}$$

и начальным $\varphi(\omega_0)$, $\omega_{\min} < \omega_0 < \omega_{\max}$ значениями. На этом свойстве непрерывности фазовой характеристики можно получить алгоритм построения частотных характеристик, если истинное значение $\varphi(\omega_0)$ лежит в пределах $(-\pi, \pi)$.

2.4. ПОЛНОТА ХАРАКТЕРИСТИК

Пусть оператор преобразования входного воздействия объекта или системы управления $f(t)$ в выходную переменную $y(t)$ представлен в форме дифференциального уравнения n -го порядка (2.1) или в форме передаточной функции (2.3).

Рассмотрим формулу (2.13) для реакции системы на воздействие $f(t)$ при нулевых предначальных условиях. Если имеет место условие

$$B(s_i) = 0,$$

где s_i — корень характеристического полинома

$$A(s_i) = 0,$$

то коэффициент C_i при экспоненте

$$\exp\{s_i t\}$$

равняется нулю при любом воздействии. По реакциям такой системы нельзя полностью выявить ее собственные свойства — в реакциях будет отсутствовать составляющая (мода), соответствующая корню s_i .

В этом случае говорят, что по рассматриваемой паре вход-выход система является *неполной*. Временные характеристики — реакции на воздействия при нулевых предначальных условиях — не отражают полностью собственных свойств системы по неполной паре вход-выход.

Операторные полиномы A и B дифференциального уравнения (2.2) неполной системы имеют нетривиальный общий делитель $(p - s_i)$, а передаточная функция (2.3) имеет *диполь* s_i . Если полиномы A и B не являются взаимно простыми, то передаточную функцию называют *вырожденной*.

$$W(s); \quad s \in C,$$

построенные при изменении аргумента s вдоль некоторого контура C на комплексной плоскости, в частности, вдоль мнимой оси $s = j\omega$; $\omega \geq 0$, т.е. частотные характеристики также отражают только полную часть системы.

Потеря части собственных свойств систем особенно существенна, если s_i — правый корень. Сокращение вырожденной передаточной функции не рекомендуется. В случае правого полюса сокращение просто недопустимо. При ненулевых начальных условиях, например, вызванных воздействиями, приложенными к другим входам системы, появляются свободные движения. Если эти начальные условия таковы, что

$$A_n(s_i) \neq 0,$$

то, как следует из (2.14), свободные движения содержат моду $\exp\{s_i t\}$.

Свойство полноты относится к модели типа M_{ysf} — системы со связями со средой — это свойство конкретной передачи.

Если даны операторные полиномы $A(p)$ и $B(p)$ или полиномы знаменателя $A(s)$ и числителя $B(s)$ передаточной функции объекта или системы управления по выбранной паре вход-выход, наличие общих делителей или диполей можно выявить несколькими способами. Во-первых, можно вычислить и непосредственно сопоставить корни полиномов A и B . Это наилучший способ, здесь выявляются и приближенные диполи передаточных функций. Во-вторых, выявить наибольший общий делитель полиномов делением их по алгоритму Евклида. В-третьих, можно исследовать *результант* полиномов — специальный определитель порядка $m + n$, построенный из коэффициентов полиномов A и B . Полиномы имеют по меньшей мере один общий корень, если их результат равен нулю.

2.5. СИСТЕМЫ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Системы дифференциальных уравнений обычно получают в результате построения математических моделей физических систем с сосредоточенными компонентами аналитическим методом.

2.5.1. Системы дифференциальных уравнений различных порядков

Пусть исходные знания об объекте управления имеют вид некоторой физической системы с сосредоточенными компонентами; это, например, может быть многоконтурная электрическая или механическая схема. На основе соответствующих законов по определенным прави-

лам записываются компонентные уравнения и уравнения связей [8, 70]. Далее эти уравнения можно привести к следующему виду:

$$\sum_{j=1}^N A_{ij}(p)x_j(t) = \sum_{r=1}^P B_{ir}(p)f_r(t); \quad i=1, \dots, N;$$

$$y_q(t) = \sum_{i=1}^N C_{qi}(p)x_i(t), \quad q=1, \dots, K. \quad (2.16)$$

Уравнения (2.16) можно записать в матричной форме

$$\begin{aligned} A(p)x(t) &= B(p)f(t); \\ y(t) &= C(p)x(t), \end{aligned} \quad (2.17)$$

где x — вектор внутренних переменных размерности N ; f и y — векторы переменных входа и выхода размерности P и K соответственно; $A(p)$, $B(p)$, $C(p)$ — полиномиальные матрицы. В уравнениях (2.17) указано, как зависят переменные выхода $y(t)$ от внутренних переменных $x(t)$. Обычно матрица C — числовая, состоит из нулей и единиц, указывающих, какие из переменных x принимаются за выходные.

Уравнения (2.16), (2.17) называют *непричинно-следственными* [8]; между внутренними переменными $x_i(t)$ нет объективных причинно-следственных отношений.

2.5.2. Системы дифференциальных уравнений в форме пространства состояний

При определенных условиях систему (2.16) можно записать в форме системы дифференциальных уравнений первого порядка, разрешенных относительно производных

$$\frac{dv_i}{dt} = a_{i1}v_1 + \dots + a_{in}v_n + b_{i1}f_1 + \dots + b_{iP}f_P; \quad i=1, \dots, n,$$

дополненной уравнениями выходов:

$$y_q(t) = c_{q1}v_1 + \dots + c_{qn}v_n; \quad q=1, \dots, K.$$

Модели в терминах вход-состояние-выход используют понятие *состояния*. Состояние динамического объекта (с памятью) — необходимая и достаточная информация для определения будущего поведения по дифференциальным уравнениям при заданных входных воздействиях независимо от того, каким путем система пришла в это состояние. Для конечномерных систем состояние представляется как n -мерный вектор

$v(t)$; при $t = 0$ вектор $v(0)$ — начальное состояние. Система дифференциальных уравнений первого порядка в так называемой нормальной форме пространства состояний (стандартизированной векторно-матричной форме Коши) записывается следующим образом:

$$\begin{aligned}\frac{dv}{dt} &= Av + Bf, v(0); \\ y &= Cv + Df,\end{aligned}\tag{2.18}$$

где f — P -мерный вектор входа; y — K -мерный вектор выхода; A — матрица состояний; B — матрица входа; C — матрица выхода; D — матрица обхода соответствующих размеров. Первую векторно-матричную строку в системе уравнений (2.18) называют уравнениями состояний, а вторую — уравнениями выхода.

Пример 4. При $n = 2$ дифференциальные уравнения (2.18) системы с одним входом и одним выходом в раскрытой форме запишутся так:

$$\frac{dv_1}{dt} = a_{11}v_1 + a_{12}v_2 + b_1f;$$

$$\frac{dv_2}{dt} = a_{21}v_1 + a_{22}v_2 + b_2f;$$

$$y = c_1v_1 + c_2v_2 + df.$$

Матрицы будут иметь следующий вид:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix};$$

$$C = [c_1 \quad c_2]; \quad D = d.$$

Если первое уравнение в системе (2.18) записать с использованием оператора дифференцирования p , то имеем:

$$(pI - A)v = Bf,$$

где I — единичная матрица. Таким образом, уравнения в форме пространства состояний являются частным случаем системы дифференциальных уравнений (2.17) с матрицей

$$A(p) = pI - A.\tag{2.19}$$

2.5.3. Дифференциальные уравнения автономных систем

Автономная система описывается однородным дифференциальным уравнением

$$A(p)y(t) = 0; \quad y(0), y'(0), \dots, y^{(n-1)}(0),$$

причем, начальные условия являются математическим отражением предыстории. Если они ненулевые, то система совершает так называемые свободные движения. В конечномерных системах свободные движения определяются полностью оператором $A(p)$ и конечным числом начальных условий независимо от того, каким путем система пришла в это состояние к моменту начала наблюдения.

Автономная система может описываться системой дифференциальных уравнений различных порядков

$$A(p)x(t) = 0, \quad x(0);$$

$$y(t) = Cx(t),$$

а также дифференциальными уравнениями в форме пространства состояний

$$\frac{dv}{dt} = Av, \quad v(0);$$

$$y = Cv.$$

2.6. ПОСТРОЕНИЕ МОДЕЛЕЙ ВХОД-ВЫХОД ПО СИСТЕМЕ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Пусть дана система дифференциальных уравнений (2.17). Построение модели в терминах «вход-выход» означает исключение внутренних переменных, что проще выполнить, если от дифференциальных уравнений перейти к системе алгебраических уравнений для изображений, приняв нулевые начальные условия:

$$\begin{aligned} A(s)x(s) &= B(s)F(s); \\ Y(s) &= Cx(s). \end{aligned} \tag{2.20}$$

2.6.1. Последовательное исключение переменных

При небольшом числе уравнений применяют метод последовательных исключений. Пусть, например, объект с одним входом f и одним выходом y имеет две внутренние переменные x_1 и x_2 :

$$\begin{aligned} A_{11}(s)X_1(s) + A_{12}(s)X_2(s) &= B_1(s)F(s); \\ A_{21}(s)X_1(s) + A_{22}(s)X_2(s) &= 0; \\ Y(s) &= X_2(s). \end{aligned} \tag{2.21}$$

Выразим вначале переменную X_1 через переменную X_2 в силу второго уравнения

$$X_1 = -\frac{A_{22}}{A_{21}}X_2.$$

Подставим это выражение в первое уравнение и найдем:

$$Y = \frac{-B_1 A_{21}}{A_{11} A_{22} - A_{12} A_{21}} F.$$

Теперь по выражению

$$W(s) = \frac{-B_1(s) A_{21}(s)}{A_{11}(s) A_{22}(s) - A_{12}(s) A_{21}(s)}$$

легко получить полиномы числителя и знаменателя передаточной функции и записать выражение для одного дифференциального уравнения. Используемые операции — перемножение и вычитание полиномов.

2.6.2. Правило Крамера

Правило Крамера (G. Cramer) удобно применять в случае, когда требуется вычислить передаточную функцию, связывающую одну из выходных переменных $y = x_q$ с одним из воздействий f_r :

$$W_{qr}(s) = \frac{\det A_{qr}(s)}{\det A(s)}, \quad (2.22)$$

где полиномиальная матрица A_{qr} получена из матрицы A заменой q -го столбца r -м столбцом матрицы B . Знаменатель передаточной функции $W_{qr}(s)$ независимо от номеров входа r и выхода q один и тот же — он равен характеристическому полиному системы

$$A(s) = \det A(s). \quad (2.23)$$

Этот способ построения моделей вход-выход по системе уравнений (2.20) сводится к вычислению определителей полиномиальных матриц.

Для примера (2.21) запишем систему в матричной форме (2.20). Матрицы имеют вид:

$$A(s) = \begin{bmatrix} A_{11}(s) & A_{12}(s) \\ A_{21}(s) & A_{22}(s) \end{bmatrix}; \quad B(s) = \begin{bmatrix} B_1(s) \\ 0 \end{bmatrix}. \quad (2.24)$$

В соответствии с правилом Крамера определяем характеристический полином и числитель передаточной функции $W_{21}(s)$ (здесь $r = 1, q = 2$):

$$A(s) = \det \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A_{11} A_{22} - A_{12} A_{21};$$

$$\det A_{21} = \begin{bmatrix} A_{11} & B_1 \\ A_{21} & 0 \end{bmatrix} = -B_1 A_{21}.$$

2.6.3. Матричный способ

Пусть имеем систему алгебраических уравнений многомерной системы, записанную для изображений переменных (2.20). В общем случае передаточная матрица системы, т. е. модель вход-выход через полиномиальные матрицы выражается так:

$$W(s) = CA^{-1}(s)B(s). \quad (2.25)$$

Здесь вычисления связаны с обращением и перемножением полиномиальных матриц. Ясно, что полиномиальная матрица системы $A(s)$ должна быть неособенной, а значит, ее определитель не равен тождественно нулю. Известно, что

$$A^{-1}(s) = \frac{A^*(s)}{\det A(s)} = \frac{A^*(s)}{A(s)},$$

где $A^*(s)$ — присоединенная матрица. Следовательно, выражение для передаточной матрицы примет вид

$$W(s) = CA^*(s)B(s)/A(s). \quad (2.26)$$

Для примера одномерной системы (2.24) характеристический полином $A(s)$ вычислен ранее. Матрица A^* , присоединенная к A , выглядит так:

$$A^* = \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}.$$

Числитель передаточной функции вычислим по формуле из (2.26):

$$(0 \quad 1) \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix} \begin{pmatrix} B_1 \\ 0 \end{pmatrix} = -B_1 A_{21}.$$

В случае одномерной системы ($K = P = 1$) полиномиальную матрицу числителя передаточной матрицы можно также вычислять как определитель следующей блочной матрицы:

$$\det \begin{bmatrix} A(s) & B(s) \\ -C & 0 \end{bmatrix}. \quad (2.27)$$

Действительно, если воспользоваться леммой Шура (I. Schur) [27], то искомый определитель раскроется так:

$$\det A(s) \det [0 + CA^{-1}(s)B(s)].$$

Учтем, что

$$A^{-1}(s) = \frac{A^*(s)}{\det A(s)}$$

и после сокращения получим выражение для числителя

$$CA^*(s)B(s).$$

Составим блочную матрицу (2.27) для рассматриваемого примера и раскроем ее определитель:

$$\det \begin{bmatrix} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & 0 \\ 0 & -1 & 0 \end{bmatrix} = -B_1 A_{21}.$$

Получим тот же результат.

2.7. ПОСТРОЕНИЕ МОДЕЛЕЙ ВХОД-ВЫХОД ПО УРАВНЕНИЯМ В ФОРМЕ ПРОСТРАНСТВА СОСТОЯНИЙ

Пусть дифференциальные уравнения объекта или системы управления записаны в форме пространства состояний:

$$\begin{aligned} \frac{dv}{dt} &= Av + Bf, v(0); \\ y &= Cv + df. \end{aligned} \quad (2.28)$$

Для простоты примем одномерный случай: переменные входа и выхода f и y являются скалярами; матрица входа B — столбец; матрица выхода C — строка; d — скаляр обхода.

2.7.1. Вычисление передаточной функции

Преобразуем уравнения (2.28) по Лапласу при нулевых начальных условиях:

$$\begin{aligned} sv(s) &= Av(s) + BF(s); \\ Y(s) &= Cv(s) + dF(s). \end{aligned} \quad (2.29)$$

Выразим решение системы алгебраических уравнений — изображение вектора состояний — в следующей форме:

$$v(s) = (sI - A)^{-1} BF(s), \quad (2.30)$$

где $(sI - A)^{-1}$ — матрица, обратная характеристической матрице $(sI - A)$ матрицы A ; I — единичная матрица. Подставим (2.30) в (2.29) и получим

$$Y(s) = W(s)F(s) = [C(sI - A)^{-1}B + d]F(s).$$

Передаточная функция W может быть записана и иначе, если учесть, что

$$(sI - A)^{-1} = (sI - A)^* / A(s), \quad (2.31)$$

где $(sI - A)^*$ — присоединенная матрица,

$$A(s) = \det(sI - A) - \quad (2.32)$$

определитель характеристической матрицы — характеристический полином системы дифференциальных уравнений (2.17). С учетом (2.31) и (2.32) передаточная функция запишется как

$$W(s) = \frac{B(s)}{A(s)} = \frac{C(sI - A)^* B + dA(s)}{A(s)}. \quad (2.33)$$

Элементами присоединенной матрицы $(sI - A)^*$ являются алгебраические дополнения элементов характеристической матрицы $(sI - A)$, т.е. полиномы. Их степени не могут превосходить $(n - 1)$. Таким образом, как видно из формулы (2.33), степень $m = \deg B$ полинома числителя передаточной функции W не может быть выше степени $n = \deg A$ характеристического полинома и равна ей только при $d \neq 0$. Это ограничивает возможности описания динамических систем в нормальной форме пространства состояний ($m \leq n$).

Имея полиномы передаточной функции (2.33), легко записать дифференциальное уравнение n -го порядка. Остается решить вопрос с перечислением начальных условий. Рассмотрим случай $d = 0$, т.е. $m < n$. Как следует из уравнения выхода системы (2.28), начальное значение выхода определяется из вектора начального состояния путем умножения на матрицу выхода:

$$y(0) = Cv(0).$$

Продифференцируем уравнение выхода по времени

$$\frac{dy}{dt} = C \frac{dv}{dt} = CAv.$$

Следовательно, начальное значение производной выхода

$$y'(0) = CAv(0).$$

Легко видеть, что

$$y''(0) = CA^2 v(0)$$

и т. д. Последнее из начальных условий для дифференциального уравнения n -го порядка определяется по формуле

$$y^{(n-1)}(0) = CA^{n-1} v(0).$$

Упорядочим начальные условия $y(0), \dots, y^{(n-1)}(0)$ в вектор $y(0)$. Тогда имеем

$$y(0) = Vy(0),$$

где V — так называемая матрица наблюдаемости:

$$V = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}.$$

Если существует V^{-1} , т.е. $\det V \neq 0$, или ранг квадратной матрицы равен ее размеру n , то можно записать зависимость начального (и текущего) состояния от значения выхода и его производных

$$v(0) = V^{-1}y(0).$$

Получение характеристического полинома A и полинома числителя передаточной функции из матриц нормальной формы пространства состояний связано с вычислением определителей полиномиальных матриц. Как следует из формулы (2.27) с учетом (2.19), полином числителя передаточной функции равен определителю следующей полиномиальной матрицы:

$$B(s) = \det \begin{bmatrix} sI - A & B \\ -C & d \end{bmatrix}. \quad (2.34)$$

С использованием формул (2.32) и (2.34) вычисление полинома знаменателя и числителя передаточной функции по матрицам нормальной формы пространства состояний сводится к вычислению определителей двух полиномиальных матриц (часто разреженных — с небольшой долей ненулевых элементов).

2.7.2. Построение временных характеристик

Преобразуем по Лапласу уравнения (2.28)

$$sv(s) - v(0) = Av(s) + BF(s)$$

получим выражение для изображения вектора состояния

$$v(s) = (sI - A)^{-1}v(0) + (sI - A)^{-1}BF(s). \quad (2.35)$$

В этой сумме первое слагаемое — свободное, а второе — вынужденное движение системы. Для получения оригинала — функции вре-

мени $v(t)$ выполняется операция обратного преобразования Лапласа. В данном случае выражение для изображения представляет собой матрицу, однако справедлива аналогия со скалярным случаем. Известно, что оригинал скалярной функции

$$L^{-1}\left\{\frac{1}{s-a}\right\}=L^{-1}\{(s-a)^{-1}\}=e^{at}$$

имеет вид экспоненты. Оказывается, что аналогичное выражение имеет место и в матричном случае, т. е.

$$L^{-1}\{s\mathbf{I}-\mathbf{A}\}^{-1}=e^{\mathbf{A}t}=\Phi(t),$$

что является матричной экспонентой, называемой матрицей перехода. Произведению изображений отвечает свертка оригиналов; это справедливо и для матриц. Поэтому вектор состояния как функция времени получается из выражения (2.35) и имеет следующий вид

$$\mathbf{v}(t)=e^{\mathbf{A}t}\mathbf{v}(0)+\int_0^te^{\mathbf{A}(t-\tau)}\mathbf{B}f(\tau)d\tau. \quad (2.36)$$

Изображение переменной выхода при нулевых начальных условиях $\mathbf{v}(0)=\mathbf{0}$ получится путем подстановки второго слагаемого выражения (2.35) во второе уравнение системы (2.28)

$$Y(s)=\left[\mathbf{C}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}+d\right]F(s).$$

Если на вход системы подается единичный импульс, т. е. $F(s)=1$, то реакция системы (импульсная переходная функция) определяется выражением:

$$y(t)=w(t)=L^{-1}\left\{\mathbf{C}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}+d\right\}=\mathbf{C}e^{\mathbf{A}t}\mathbf{B}+d\delta(t). \quad (2.37)$$

Сопоставляя полученную формулу с выражением для передаточной функции (2.33), замечаем, что

$$\Phi(t)=e^{\mathbf{A}t}=L^{-1}\left\{(s\mathbf{I}-\mathbf{A})^{-1}\right\}.$$

Отсюда следует один из способов получения матрицы перехода путем обращения по Лапласу матрицы $(s\mathbf{I}-\mathbf{A})^{-1}$.

Пример 5. Пусть матрица состояний нормальной формы имеет вид

$$\mathbf{A}=\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}.$$

Характеристическая матрица запишется как

$$(sI - A) = \begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix},$$

а ее обращение дает

$$(sI - A)^{-1} = \frac{(sI - A)^*}{A(s)} = \frac{1}{s^2 + 3s + 2} \begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix}.$$

Применение обратного преобразования Лапласа к каждому элементу полученной матрицы позволяет получить выражение для матрицы перехода

$$\Phi(t) = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}.$$

Отсюда по формуле (2.37) легко находится функция веса, если известны матрицы выхода $C = (2 \ 0)$ и входа $B = (0 \ 1)^T$:

$$w(t) = (2 \ 0)\Phi(t)(0 \ 1)^T = 2e^{-t} - 2e^{-2t}.$$

Переходная функция $h(t)$ является интегралом от функции веса

$$h(t) = C \int_0^t e^{A\tau} d\tau B + d1(t) = CA^{-1}(e^{At} - I)B + d1(t).$$

Построение временных характеристик по нормальной форме пространства состояний связано с вычислением матричного экспоненциала — матрицы перехода $\Phi(t)$. Выше приведены примеры получения аналитических выражений для элементов $\Phi(t)$ путем обратного преобразования Лапласа. Аналитическое выражение для $\Phi(t)$ может быть получено по формуле Лагранжа — Сильвестра (J. L. Lagrange — J. Sylvester):

$$\Phi(t) = e^{At} = \sum_{i=1}^n \frac{(A - s_1 I) \dots (A - s_{i-1} I)(A - s_{i+1} I) \dots (A - s_n I)}{(s_i - s_1) \dots (s_i - s_{i-1})(s_i - s_{i+1}) \dots (s_i - s_n)} e^{s_i t} \quad (2.38)$$

(случай простых собственных значений s_i).

Другим способом численного получения значений матрицы перехода при фиксированных значениях t является разложение матричного экспоненциала в степенной ряд

$$e^{At} = I + At + \frac{(At)^2}{2!} + \dots + \frac{(At)^M}{M!} + \dots$$

2.8. МОДЕЛИ СИСТЕМ УПРАВЛЕНИЯ С РАСКРЫТОЙ ПРИЧИННО-СЛЕДСТВЕННОЙ СТРУКТУРОЙ

Под *структурой систем управления* понимают причинно-следственную связь между элементами направленного действия. Понятия «система» и «структура» являются близкими по смыслу. Наиболее общие определения понятий системы и структуры строятся как отношения на множествах. Математически — это *графы*. Графы являются универсальным средством описания структур систем. При небольшом числе элементов и связей весьма наглядны *диаграммы графов*, т.е. их геометрические образы. В зависимости от элементов множеств рассматриваются различные типы графов. Приведенные в 1.4 диаграммы, иллюстрирующие принципы управления, отражают типовые структуры причинно-следственных отношений основных элементов систем управления и, по существу, представляют собой *ориентированные графы*. Там же использовались и такие термины теории орграфов, как *контуры* и *пути*. Электрическая и механическая схемы, изображенные на рис. 1.20, также являются примерами графов, только неориентированных. Имея в виду структуру связей элементов, иногда говорят о *топологии* (топографии) системы. Даже без конкретизации вершин и дуг, т.е. только по топологии, можно сделать ряд важнейших выводов о свойствах системы, которые сохраняются при дальнейшем раскрытии неопределенности — уточнении структур операторов и конкретизации значений параметров.

В зависимости от подхода к моделированию, от конкретного содержания элементов исходного множества и элементов отношения модели с раскрытой структурой могут быть представлены *структурными схемами*, *сигнальными графами*, *системами дифференциальных уравнений в причинно-следственной форме* и некоторыми другими формами.

2.8.1. Структурные схемы (С-графы)

Структурная схема представляет собой причинно-следственную связь *звеньев*. Линейное звено (рис. 2.6, а) в общем случае имеет любое число входов; оно преобразует сумму входов в единственную переменную выхода по некоторому оператору w_i (рис. 2.6, б)

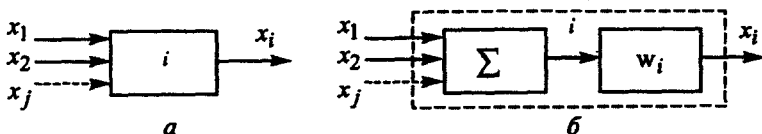


Рис. 2.6. Линейное звено

$$x_i = w_i \left(\sum_j x_j \right) = \sum_j w_i(x_j).$$

В частном случае оператора тождественного преобразования звено выступает как сумматор.

Структурная схема является ориентированным графом и состоит из множества вершин $W = \{w_1, \dots, w_N\}$ и множества дуг $X = \{(w_i, w_j)\}$ — упорядоченных пар вершин [25]. Дугам графа соответствуют переменные x_i ; $i = 1, \dots, N$, а вершинам — звенья. Для того чтобы отличать рассматриваемый граф от сигнальных графов других типов, назовем его C -графом. На языке теории бинарных отношений C -граф определяется как пара множеств

$$C = \langle W, X \rangle,$$

а структурная схема — геометрический образ — называется также диаграммой графа (рис. 2.7). Вершина C -графа — звено общего вида, по определению суммирует переменные заходящих дуг. Это позволяет отказаться от специального элемента суммирования, что отличает C -графы от классических структурных схем.

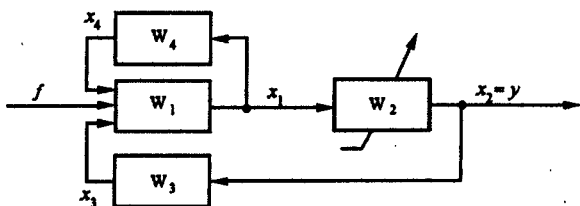


Рис. 2.7. Структурная схема (C -граф)

Дуга C -графа — элемент (w_i, w_j) отношения X — задает причинно-следственную связь между двумя звеньями, причем, выход j -го звена является входом i -го. Дуге (w_i, w_j) соответствует переменная x_j .

Теоретико-множественное описание систем дает естественный способ ввода и редактирования моделей систем управления как последовательного раскрытия неопределенности [21]. Для этого модели упорядочиваются по рангам неопределенности $R = 0, 1, 2, 3$.

Множество W звеньев задает модель нулевого ранга $M_0(0)$. Для примера C -графа, диаграмма которого изображена на рис. 2.7, множество перечисляется так:

$$W = \{w_1, w_2, w_3, w_4\}.$$

В случае однотипных звеньев можно ограничиться заданием числа вершин графа (звеньев), т.е. мощности множества $|W| = N = 4$.

Дополнение модели $M_s(0)$ множеством X дает модель первого ранга $M_s(1)$ — это топология (топография) системы. Для примера C -графа (см. рис. 2.7) множество перечисляется так:

$$X = \{(1,3), (1,4), (2,1), (3,2), (4,1)\}.$$

В перечислении приведены только индексы (номера) звеньев.

Дальнейшее раскрытие неопределенности достигается при задании структур операторов вершин. Для рассматриваемого класса систем передаточные функции являются отношениями полиномов:

$$W_i(s) = \frac{B_i(s)}{A_i(s)}.$$

Задание их структур сводится к указанию степеней m_i и n_i полиномов B_i и A_i .

Когда для всех звеньев заданы структуры операторов, образуется модель системы структурного ранга $M_s(2)$.

Пусть для рассматриваемого примера системы (см. рис. 2.7) передаточные функции звеньев имеют вид:

$$W_1(s) = k_1; W_2(s) = k_2 / (1 + T_2 s)^2; W_3(s) = -1; W_4(s) = -\tau_4 s / (1 + T_4 s).$$

Информацию о структурах операторов можно закодировать массивами степеней полиномов числителей и знаменателей передаточных функций:

$$\{0, 0, 0, 1\} \text{ и } \{0, 2, 0, 1\}.$$

Результатом конкретизации значений всех коэффициентов полиномов является полностью определенная модель третьего, параметрического ранга $M_s(3)$.

Выше изложено описание собственно системы (автономной системы). Для описания связей системы со средой следует указать звено, ко входу которого подается воздействие, и звено, выход которого является выходом системы. На примере C -графа (см. рис. 2.7) номер входного звена $r = 1$, а выходного — $q = 2$. В результате оказывается определенной модель системы со связями со средой $M_{TSF}(3)$.

При изучении влияния вариаций звеньев на характеристики системы указывается варьируемое звено. На рис. 2.7 им является звено w_2 .

2.8.2. Сигнальные графы

Сигнальный граф или граф Мэсона (S. J. Mason) является одной из удобных в теории и расчетной практике форм представления моделей систем управления.

Модель системы в форме сигнального графа определяется как бинарное отношение W на множестве переменных $X = \{x_1, \dots, x_N\}$:

$$G = \langle X, W \rangle.$$

Элементом отношения $W = \{(x_i, x_j)\}$ ставятся в соответствие операторы преобразования переменных. На диаграммах сигнальных графов переменным отвечают вершины, где суммируются сигналы заходящих дуг, а элементам отношения — дуги.

Способы задания моделей различных рангов в форме сигнальных графов — те же, что и для C -графов.

На рис. 2.8 изображена диаграмма сигнального графа — модель топологического ранга, несущая ту же информацию о системе, что и структурная схема (см. рис. 2.7).

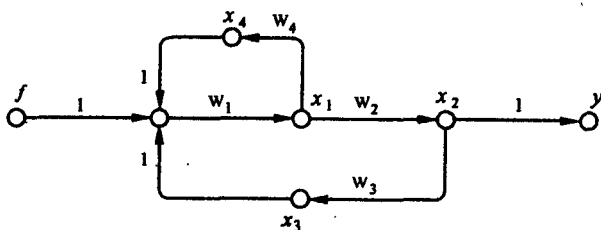


Рис. 2.8. Диаграмма сигнального графа

В заключение подчеркнем, что формы представления моделей и способы их отображения могут быть различными — символьными или алгебраическими (уравнения, матрицы), геометрическими или топологическими (диаграммы графов). Однако информация о моделях различных рангов R последовательно раскрывается описанием множеств, задающих: $R = 0$ — состав элементов; $R = 1$ — топологию причинно-следственных связей между ними; $R = 2$ — структуры операторов; $R = 3$ — параметры.

Теоретико-множественное представление структур систем в форме графов обеспечивает формализацию описания моделей, упрощает кодирование их графических образов, а также разработку алгоритмов анализа систем. На базе C -графов построено описание систем управления в программе *CLASSiC* для персональных компьютеров [4], а на базе графа Мизона — описание систем в программе *АРДИС* для СМ ЭВМ [65].

2.8.3. Причинно-следственные системы дифференциальных уравнений

Пусть в каждой строке $i = 1, \dots, N$ системы уравнений (2.16) присутствует член с переменной x_r , которая является следствием остальных

ных переменных. Диагональный оператор назовем собственным оператором звена, описываемого i -м уравнением. Недиагональные полиномы $A_{ij}(p)$; $i \neq j$, и $B_{ir}(p)$ являются операторами воздействий на входе i -го звена. Таким образом, причинно-следственная система уравнений [8] отличается от непричинно-следственной тем, что конкретизируется порядок перечисления уравнений, причем, $A_{ii}(p) \neq 0$; $i = 1, \dots, N$.

Если система автономна, то $B_{ir}(p) \equiv 0$; $i = 1, \dots, N$, $r = 1, \dots, P$. Заметим, что пары индексов ненулевых полиномов $A_{ij}(p)$; $i \neq j$ в точности образуют элементы бинарного отношения модели в форме C -графа.

2.9. ПРЕОБРАЗОВАНИЕ ФОРМ ПРЕДСТАВЛЕНИЯ МОДЕЛЕЙ С РАСКРЫТОЙ СТРУКТУРОЙ

2.9.1. Запись дифференциальных уравнений по графу

Для получения по C -графу модели в форме системы дифференциальных уравнений вначале записываются алгебраические уравнения для изображений переменных вида (2.20). Далее, формальной заменой символа комплексного аргумента s на оператор дифференцирования по времени $p \equiv d/dt$, получаются дифференциальные уравнения.

Полиномиальные матрицы системы (2.17) непосредственно заполняются по C -графу. Диагональными элементами полиномиальной матрицы $A(s)$ являются полиномы знаменателей передаточных функций звеньев, а ненулевыми недиагональными элементами — полиномы числителей, взятые с обратными знаками, т. е.

$$A_{ij}(s) = \begin{cases} A_i(s), & i = j; \\ -B_i(s), & (i, j) \in X, i \neq j; \\ 0, & (i, j) \in X, i \neq j. \end{cases}$$

Матрица $B(s)$ имеет отличные от нуля элементы $B_i(s)$ в r -х строках, где r — номера входных блоков. Матрица C состоит из единиц и нулей, причем единицы располагаются в столбцах с номерами q выходных блоков.

Полученная система дифференциальных уравнений имеет причинно-следственную форму.

Пример 6. Получим вид матриц A и B для примера C -графа (см. рис. 2.7). Система уравнений для изображений переменных запишется так:

$$x_1 = W_1(x_3 + x_4 + f);$$

$$x_2 = W_2 x_1;$$

$$x_3 = W_3 x_2;$$

$$x_4 = W_4 x_1.$$

Учитывая выражения для передаточных функций звеньев, имеем после приведения к общему знаменателю и переноса части членов в левую часть:

$$A_1 x_1 - B_1 x_3 - B_1 x_4 = B_1 f;$$

$$-B_2 x_1 + A_2 x_2 = 0;$$

$$-B_3 x_2 + A_3 x_3 = 0;$$

$$-B_4 x_1 + A_4 x_4 = 0.$$

Введем вектор переменных $x = (x_1, x_2, x_3, x_4)^T$ и запишем уравнение в компактной форме (2.18). Соответствующие матрицы имеют вид:

$$A = \begin{bmatrix} A_1 & 0 & -B_1 & -B_1 \\ -B_2 & A_2 & 0 & 0 \\ 0 & -B_3 & A_3 & 0 \\ -B_4 & 0 & 0 & A_4 \end{bmatrix}; \quad B = \begin{bmatrix} B_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Обратим внимание на то, что структура матрицы задает топологию системы — модель $M(1)$. Действительно, пары номеров строк и столбцов (i, j) ненулевых элементов вне главной диагонали в точности образуют множество X .

2.9.2. Построение графов по дифференциальным уравнениям

Для построения ориентированного графа система дифференциальных уравнений (2.16) предварительно преобразуется по Лапласу, в результате чего получается система алгебраических уравнений для изображений переменных. При нулевых начальных условиях эта операция сводится к формальной замене символа оператора дифференцирования $p \equiv d/dt$ на символ комплексного аргумента s .

Алгоритм построения графа Мэзона по системе уравнений складывается из следующих операций [8]:

- выбирается такая последовательность уравнений, в которой каждое i -е уравнение содержит переменную $x_i(s)$;
- каждое i -е уравнение делится на полином $A_{ii}(s) \neq 0$;
- слагаемые с переменными x_j ; $j \neq i$, вида $(A_{ij}/A_{ii})x_j$ переносятся в правые части уравнений.

Полученная причинно-следственная форма записи уравнений по сути является графом Мэзона.

Для любой системы уравнений (2.16), (2.17) с невырожденной матрицей существует такая последовательность уравнений, когда все элементы A_{ii} главной диагонали матрицы A отличны от нуля [8]. Следовательно, для каждой системы уравнений с отличным от нуля определителем $\det A(s)$ можно построить ориентированный граф. Вместе с тем, такая последовательность может быть неединственной, т.е. переход от уравнений к графам в общем случае не является однозначным. Выбор одного из вариантов графа по существу означает дополнение динамической модели информацией о направленных взаимодействиях элементов, т.е. переход к моделям, характерным для теории управления.

2.10. ХАРАКТЕРИСТИКИ СИСТЕМ С ТИПОВЫМИ СТРУКТУРАМИ

Системы с типовой структурой образуются последовательным (рис. 2.9, а), параллельным (рис. 2.9, б) соединениями звеньев или соединением с обратной связью (рис. 2.9, в).

Выявление свойств типовых систем в целом связано с построением эквивалентных систем со свернутой структурой (рис. 2.9, г). Эквивалентные системы в терминах вход-выход могут быть представлены в форме дифференциального уравнения

$$A_3(p)y(t) = B_3(p)f(t), \quad (2.39)$$

передаточной функции

$$W_3(s) = \frac{Y(s)}{F(s)} = \frac{B_3(s)}{A_3(s)},$$

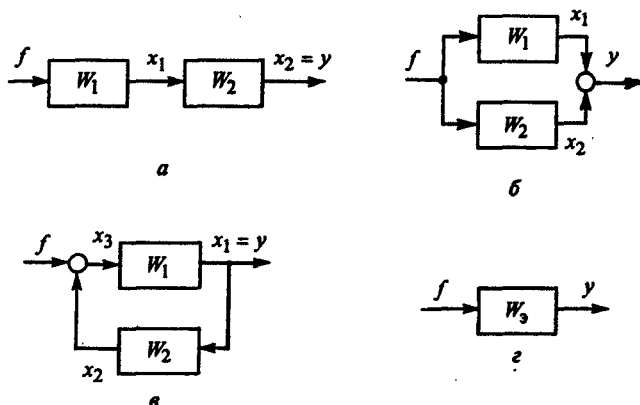


Рис. 2.9. Системы с типовой структурой

временной характеристики

$$w_3(t) = L^{-1} \{W_3(s)\}$$

или частотной характеристики

$$W_3(j\omega) = R_3(\omega) e^{j\varphi_3(\omega)} = P_3(\omega) + jQ_3(\omega).$$

2.10.1. Характеристики последовательного соединения

Дифференциальные уравнения системы, образованной последовательным соединением звеньев, запишутся как

$$A_1(p)x_1(t) = B_1(p)f(t);$$

$$A_2(p)x_2(t) = B_2(p)x_1(t);$$

$$y(t) = x_2(t).$$

В результате исключения переменных x_1 и x_2 получим операторные полиномы уравнения (2.39):

$$A_3(p) = A_1(p)A_2(p); \quad B_3(p) = B_1(p)B_2(p).$$

Одновременно запишем передаточную функцию эквивалентного звена:

$$W_3(s) = \frac{B_1(s)B_2(s)}{A_1(s)A_2(s)} = W_1(s)W_2(s). \quad (2.40)$$

Временную характеристику — импульсную переходную функцию, получаем обратным преобразованием Лапласа передаточной функции (2.40):

$$w_3(t) = L^{-1} \{W_1(s)W_2(s)\}.$$

Амплитудно-частотная характеристика равна произведению

$$R_3(\omega) = R_1(\omega)R_2(\omega),$$

и фазо-частотная характеристика — сумме

$$\varphi_3(\omega) = \varphi_1(\omega) + \varphi_2(\omega)$$

соответствующих характеристик последовательно соединенных звеньев. ЛАЧХ системы получается в виде суммы:

$$L_3(\omega) = L_1(\omega) + L_2(\omega).$$

На рис. 2.10 изображен пример графического построения асимптотической ЛАЧХ системы, образованной последовательным соединением интегрирующего звена W_1 и апериодического звена первого порядка W_2 .

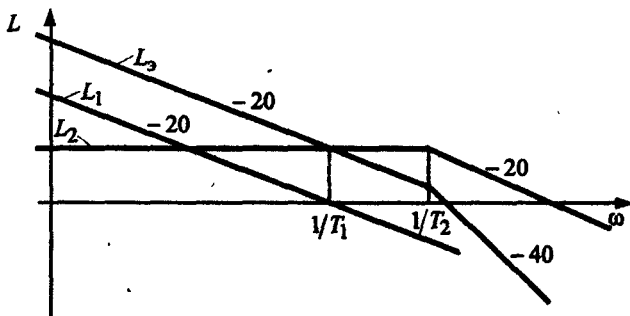


Рис. 2.10. Пример построения асимптотической ЛАЧХ последовательного соединения

Если передаточная функция $W_3(s)$ имеет простой диполь

$$B_3(s_i) = A_3(s_i) = 0,$$

то соответствующий коэффициент разложения (2.11) на сумму простейших дробей равен нулю:

$$C_i = \frac{B_3(s_i)}{A'_3(s_i)} = 0.$$

Это значит, что реакция системы на единичную δ -функцию при нулевых предначальных условиях будет неполной — она не содержит слагаемого, соответствующего полюсу s_i . Диполь $W_3(s)$ может иметь место, если имеют диполь $W_1(s)$ или $W_2(s)$, а также если

$$B_1(s_i) = A_2(s_i) = 0$$

или

$$B_2(s_i) = A_1(s_i) = 0,$$

т. е. нуль одного звена компенсирует полюс другого.

Если передаточная функция имеет диполь, то частотные характеристики, как и временные, не полностью представляют свойства системы.

2.10.2. Характеристики параллельного соединения

Дифференциальные уравнения системы, образованной параллельным соединением звеньев (см. рис. 2.9, б), запишутся как

$$A_1(p)x_1(t) = B_1(p)f(t);$$

$$A_2(p)x_2(t) = B_2(p)f(t);$$

$$x_3(t) = x_1(t) + x_2(t);$$

$$y(t) = x_3(t).$$

В результате исключения переменных x_i получим операторные полиномы эквивалентного уравнения (2.39)

$$A_3(p) = A_1(p)A_2(p);$$

$$B_3(p) = B_1(p)A_2(p) + A_1(p)B_2(p).$$

Передаточная функция эквивалентного звена

$$W_3(s) = \frac{B_1(s)A_2(s) + A_1(s)B_2(s)}{A_1(s)A_2(s)} = W_1(s) + W_2(s) \quad (2.41)$$

равна сумме передаточных функций звеньев.

Временная характеристика системы является суммой временных характеристик звеньев

$$w_3(t) = w_1(t) + w_2(t).$$

При параллельном соединении звеньев легко получить вещественную $P_3(\omega)$ и мнимую $Q_3(\omega)$ частотные характеристики эквивалентного звена

$$P_3(\omega) = P_1(\omega) + P_2(\omega);$$

$$Q_3(\omega) = Q_1(\omega) + Q_2(\omega).$$

Диполь передаточной функции $W_3(s)$, как следует из (2.41), получается, если:

- одна из передаточных функций звеньев имеет диполь;
- звенья имеют одинаковые полюсы

$$A_1(s_i) = A_2(s_i) = 0.$$

2.10.3. Характеристики соединения с обратной связью

Дифференциальные уравнения типового соединения с обратной связью (см. рис. 2.9, в) запишутся как

$$A_1(p)x_1(t) = B_1(p)x_3(t);$$

$$A_2(p)x_2(t) = B_2(p)x_1(t);$$

$$x_3(t) = f(t) \pm x_2(t);$$

$$y(t) = x_1(t),$$

где знак «минус» соответствует отрицательной обратной связи, а знак «плюс» — положительной.

Исключение внутренних переменных дает операторные полиномы дифференциального уравнения эквивалентного звена

$$\begin{aligned} A_3(p) &= A_1(p)A_2(p) \pm B_1(p)B_2(p); \\ B_3(p) &= B_1(p)A_2(p). \end{aligned} \quad (2.42)$$

Передаточная функция эквивалентного звена выражается так

$$W_3(s) = \frac{B_1(s)A_2(s)}{A_1(s)A_2(s) \pm B_1(s)B_2(s)} = \frac{W_1(s)}{1 \pm W_1(s)W_2(s)}. \quad (2.43)$$

Если звенья образуют контур положительной обратной связи, то в формулах (2.42), (2.43) используется знак «минус».

Временная характеристика системы с обратной связью $w_3(t)$ сложным образом зависит от $w_1(t)$ и $w_2(t)$, поэтому ее удобнее получать обратным преобразованием Лапласа эквивалентной передаточной функции:

$$w_3(t) = L^{-1}\{W_3(s)\}.$$

Комплексная частотная характеристика системы с обратной связью

$$W_3(j\omega) = \frac{W_1(j\omega)}{1 \pm W_1(j\omega)W_2(j\omega)} \quad (2.44)$$

также сложным образом зависит от частотных характеристик звеньев.

Свойства системы с обратной связью определяются усилением разомкнутого контура с передаточной функцией

$$W_p(s) = W_1(s)W_2(s)$$

на различных частотах. Если усиление контура мало, то можно пренебречь обратной связью. Действительно, по виду выражения (2.44) можно заключить, что на частотах, где выполняется условие

$$|W_p(j\omega)| = |W_1(j\omega)W_2(j\omega)| \ll 1,$$

имеет место приближенное соотношение

$$W_3(j\omega) \approx W_1(j\omega).$$

Практически усиление контура считается малым, если

$$L_p(\omega) = 20 \lg |W_p(j\omega)| < -16 \dots -20 \text{ дБ}.$$

С другой стороны, на частотах, где выполняется условие

$$|W_p(j\omega)| \gg 1,$$

имеет место другое приближенное соотношение

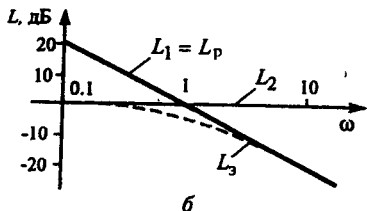
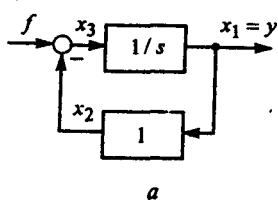


Рис. 2.11. Пример системы с обратной связью (а) и построение эквивалентной ЛАЧХ (б)

$$W_3(j\omega) \approx 1/W_2(j\omega).$$

Система в целом имеет частотную характеристику, близкую к обратной частотной характеристике звена обратной связи. Практически усиление велико, если

$$L_p(\omega) > 16...20 \text{ дБ.}$$

На остальных частотах, где

$$-16 \text{ дБ} < L_p(\omega) < 16 \text{ дБ,}$$

необходимо пользоваться точной формулой (2.44) или специальными номограммами замыкания [13, 64, 75, 90].

Рассмотрим пример системы, образованной интегрирующим звеном, охваченным единичной отрицательной обратной связью (рис. 2.11, а). На рис. 2.11, б изображены логарифмические амплитудно-частотные характеристики L_1 и L_2 этих звеньев. На частотах $\omega < 0,1 \text{ с}^{-1}$ усиление контура превышает 20 дБ. Следовательно, амплитудно-частотная характеристика замкнутой системы на этих частотах определяется только свойствами звена обратной связи, т.е. замкнутая система на низких частотах с большой степенью приближения ведет себя как безынерционное звено с единичным усилением. Напротив, на частотах $\omega > 10 \text{ с}^{-1}$ усиление контура ниже -20 дБ. Здесь контур практически разомкнут — замкнутая система ведет себя как интегрирующее звено. На рис. 2.11, б пунктирной линией изображена логарифмическая амплитудно-частотная характеристика замкнутой системы, эквивалентной аperiodическому звену первого порядка

$$W_3(s) = \frac{1/s}{1 + 1/s} = \frac{1}{s+1}.$$

Передаточная функция системы с обратной связью $W_3(s)$ имеет диполь, т.е. неполна, если передаточная функция одного из звеньев имеет диполь или нуль звена в прямом пути равен полюсу звена обратной связи, т.е. наибольший общий делитель

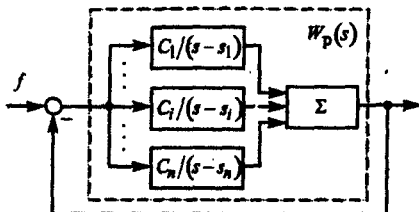


Рис. 2.12. Иллюстрация неполноты передаточной функции контура

$$A_d = \text{НОД}(B_1, A_2)$$

— нетривиальный. При этом никакими воздействиями при нулевых предначальных условиях нельзя полностью обнаружить собственные свойства замкнутой системы. Диполь $W_p(s)$ означает также наличие дipoля у передаточной функции разомкнутого контура $W_p(s)$. Это означает, что характеристический полином замкнутой системы A_d (2.31) имеет сомножитель A_d , т. е. замыкание системы перемещает не все корни — часть корней характеристического полинома замкнутой системы совпадает с корнями полинома разомкнутой системы.

На комплексной частоте нуля передаточной функции W_p усиление контура равно нулю, т. е. контур как бы разомкнут на соответствующей комплексной частоте. Если W_p имеет такой полюс, то в разложении W_p на сумму простейших дробей соответствующий коэффициент C_i равен нулю.

На рис. 2.12 изображена структурная схема системы с единичной обратной связью, где звено в прямой цепи

$$W_1(s) = W_p(s) = \sum_{i=1}^n \frac{C_i}{s-s_i}$$

представлено как параллельное соединение простейших звеньев. При наличии дipoля $C_i = 0$ и соответствующее звено оказывается «висячим», т. е. не охватывается обратной связью.

2.11. ПОСТРОЕНИЕ МОДЕЛЕЙ ВХОД-ВЫХОД СИСТЕМ С РАСКРЫТОЙ СТРУКТУРОЙ

Модели с раскрытой структурой в форме графов, структурных схем или дифференциальных уравнений в причинно-следственной форме являются развернутыми моделями объектов и систем управления. Задачи анализа обычно требуют исследования непосредственных связей выходов со входами, т. е. рассмотрения свернутых моде-

лей в терминах вход-выход: передаточных функций; частотных или временных характеристик.

Для построения моделей «вход-выход» — характеристик систем с раскрытой причинно-следственной структурой — можно предварительно записать системы уравнений и воспользоваться способами, описанными в § 2.6. Передаточные функции структурных схем и графов можно получить и непосредственно. Для этого применяют последовательное эквивалентное преобразование графов или формулу Мэсона.

2.11.1. Последовательное применение правил эквивалентного преобразования графов

Основными правилами эквивалентных преобразований графов и структурных схем являются полученные в § 2.10 правила получения эквивалентных передаточных функций типовых соединений. Они соответствуют правилам исключения переменных в системах уравнений (см. п. 2.6.1).

Пример 7. Рассмотрим пример получения передаточной функции $\Phi_{31}(s)$ для структурной схемы, изображенной на рис. 2.13. Входом системы является вход звена $r = 1$, выходом — выход звена $q = 5$.

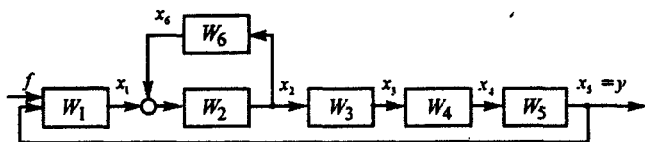


Рис. 2.13. Пример структурной схемы

Наиболее рациональная последовательность преобразований здесь следующая. Звенья W_2 и W_6 образуют контур; их можно заменить одним звеном с передаточной функцией

$$W_2 / (1 - W_2 W_6).$$

В полученной структурной схеме звенья в прямом пути соединены последовательно; их можно заменить одним звеном с передаточной функцией:

$$W_1 W_2 W_3 W_4 W_5 / (1 - W_2 W_6).$$

В результате получена типовая структура с обратной связью, передаточная функция которой равна

$$\Phi_{31} = \frac{W_1 W_2 W_3 W_4 W_5}{1 - W_2 W_6 - W_1 W_2 W_3 W_4 W_5}.$$

Ясно, что последовательность преобразований, приводящая к цели, в общем случае неединственна. С другой стороны, последовательность

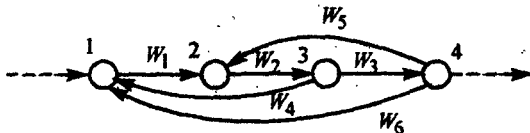


Рис. 2.14. Пример графа, не имеющего типовых подграфов

преобразований, использующая только три правила для типовых структур, существует не всегда. На рис. 2.14 изображен пример такого графа.

Необходимо ввести дополнительные правила — переноса вершин сѐма (рис. 2.15) и суммирования (рис. 2.16) сигналов.

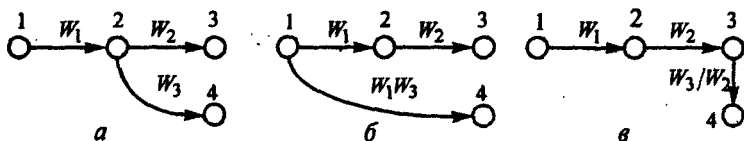


Рис. 2.15. Иллюстрация правила переноса вершины сѐма

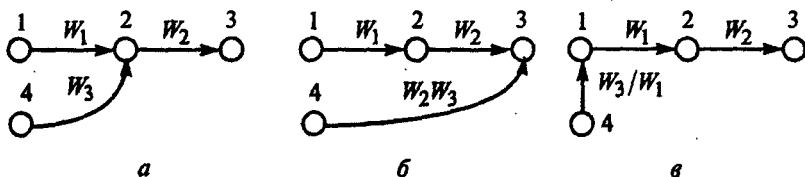


Рис. 2.16. Иллюстрация правила переноса вершины суммирования

Покажем, что дополнительные правила позволяют найти последовательность преобразований графа (см. рис. 2.14). Перенесем вершину сѐма, т. е. вместо дуги (1, 3) введем еще одну дугу (1, 4) с передаточной функцией W_4/W_3 . В результате получим граф, преобразования которого по известным трем правилам позволяют найти искомую передаточную функцию:

$$\Phi_{41} = \frac{W_1 W_2 W_3}{1 - W_2 W_3 W_5 - W_1 W_2 W_4 - W_1 W_2 W_3 W_6}.$$

2.11.2. Формула Мэсона

Пусть исходная модель представлена в форме структурной схемы (С-графа) или сигнального графа Мэсона. Для вычисления передаточной функции можно:

- записать систему уравнений (2.16) и воспользоваться формулой (2.22);

- провести последовательную топологическую редукцию графов по правилам эквивалентных преобразований;

- применить топологическую формулу Мэсона.

Формула Мэсона для получения передаточных функций сигнальных графов является топологическим аналогом правила Крамера.

Прежде всего введем несколько понятий из теории сигнальных графов. *Путь* в графе называется последовательность вершин и дуг, в которой каждый элемент встречается только раз. *Контур* — замкнутый путь. Контуров называются *некасающимися*, если они не имеют общих вершин. Передаточные функции путей и контуров равны произведению передаточных функций образующих их дуг.

Передаточная функция графа от вершины r (вход) до вершины q (выход) равна

$$\Phi_{qr}(s) = \frac{\sum P_{qr}^p(s) \Delta_{qr}^p(s)}{\Delta(s)}, \quad (2.45)$$

где $\Delta(s)$ — определитель графа:

$$\Delta(s) = 1 - \sum_k K_k(s) + \sum_{k,l} K_k(s) K_l(s) - \sum_{k,l,m} K_k(s) K_l(s) K_m(s) + \dots, \quad (2.46)$$

причем $K_k(s)$ — передаточные функции контуров графа; $K_k(s) K_l(s)$ — произведения передаточных функций некасающихся пар контуров; $K_k(s) K_l(s) K_m(s)$ — произведения передаточных функций попарно некасающихся троек контуров; $P_{qr}^p(s)$ — передаточная функция p -го пути в графе от вершины r до вершины q ; $\Delta_{qr}^p(s)$ — минор p -го пути, равный определителю подграфа, полученного удалением из исходного графа p -го пути.

Пример 8. Рассмотрим диаграмму графа, изображенную на рис. 2.8. Граф имеет два касающихся контура с передаточными функциями:

$$K_1(s) = W_1(s) W_2(s) W_3(s);$$

$$K_2(s) = W_1(s) W_4(s).$$

Определитель графа равен:

$$\Delta(s) = 1 - K_1(s) - K_2(s).$$

Между вершиной входа f и вершиной выхода y имеется один путь с передаточной функцией

$$P_{yf}^1(s) = W_1(s) W_2(s).$$

После удаления этого пути не остается ни одного контура, поэтому минор пути равен единице

$$\Delta_{yf}^1(s) = 1.$$

Таким образом, передаточная функция графа между вершинами f и u в соответствии с формулой (2.45) равна:

$$\Phi_{yf}(s) = \frac{W_1(s)W_2(s)}{1 - W_1(s)W_2(s)W_3(s) - W_1(s)W_4(s)}.$$

Если в эту формулу подставить дробно-рациональные выражения для передаточных функций (2.28), то получим:

$$\Phi_{yf}(s) = \frac{B_{yf}(s)}{A(s)} = \frac{k_1 k_2 (1 + T_4 s)}{(1 + T_2 s)^2 (1 + T_4 s) + k_1 k_2 (1 + T_4 s) + k_1 \tau_4 s (1 + T_2 s)^2}.$$

Знаменатель этого выражения $A(s)$ представляет собой характеристический полином (ХП) системы.

ХП системы, в которой разомкнуты все контуры, равен произведению ХП (знаменателей передаточных функций) всех дуг

$$A_0(s) = \prod_{\gamma=1}^N A_{\gamma}(s).$$

Можно показать [22], что определитель графа равен отношению ХП $A(s)$ и $A_0(s)$ соответственно исходной системы и системы, в которой все контуры разомкнуты:

$$\Delta(s) = A(s) / A_0(s).$$

Очевидно, для бесконтурного графа имеет место: $A(s) \equiv A_0(s)$ и $\Delta(s) = A_0(s)/A_0(s) \equiv 1$.

2.12. ПОСТРОЕНИЕ МОДЕЛЕЙ С РАСКРЫТОЙ СТРУКТУРОЙ ПО МОДЕЛЯМ ВХОД-ВЫХОД

Обратный переход от моделей вход-выход к моделям с раскрытой структурой не является однозначным. Действительно, одной и той же передаточной функции могут соответствовать сколько угодно систем дифференциальных уравнений различных порядков (2.17), уравнений в форме пространства состояний (2.18), а также сигнальных графов или структурных схем.

2.12.1. Построение модели в форме пространства состояний по дифференциальному уравнению n -го порядка

Для одномерной системы легко записать матрицы канонических форм пространства состояний по передаточной функции. Построение развернутых моделей для многомерных объектов, описанных, например, передаточными матрицами, представляет более сложную задачу так называемой минимальной реализации [37].

Переход от одного дифференциального уравнения n -го порядка (2.1) к системе уравнений в нормальной форме пространства состояний (2.18) неоднозначен из-за произвола в выборе переменных состояний. Действительно, невырожденное преобразование вектора переменных состояния

$$v = Tw; \quad \det T \neq 0, \quad (2.46)$$

где w — новый вектор состояния, не изменяет оператора преобразования вход-выход. Той же самой остается и передаточная функция. Покажем это.

Подставляя (2.46) в (2.18), получим (при коэффициенте обхода $d=0$)

$$T \frac{dw}{dt} = ATw + Bf.$$

$$y = CTw.$$

Поскольку матрица преобразования T невырожденная, существует обратная матрица T^{-1} . С учетом этого

$$\frac{dw}{dt} = T^{-1}ATw + T^{-1}Bf.$$

Введя новые обозначения $\tilde{A} = T^{-1}AT$; $\tilde{B} = T^{-1}B$; $\tilde{C} = CT$, получим

$$\begin{aligned} \frac{dw}{dt} &= \tilde{A}w + \tilde{B}f; \\ y &= \tilde{C}w. \end{aligned} \quad (2.47)$$

Системе уравнений в нормальной форме (2.47) соответствует та же передаточная функция, что и системе (2.18). По формуле, полученной в предыдущем параграфе, имеем

$$\tilde{W}(s) = \tilde{C}(sI - \tilde{A})^{-1} \tilde{B}.$$

Выразим новые матрицы $\tilde{A}, \tilde{B}, \tilde{C}$ через исходные и проведем преобразования

$$\begin{aligned} \tilde{W}(s) &= CT(sI - T^{-1}AT)^{-1} T^{-1}B = CT \left[T^{-1}(sI - A)^{-1} T \right] T^{-1}B = \\ &= C(sI - A)^{-1} B = W(s). \end{aligned}$$

Чаще всего выбирают следующие связи между переменными состояниями

$$v_2 = v_1', \quad v_3 = v_2', \dots, v_n = v_{n-1}'.$$

Это сразу определяет первые $n-1$ уравнений состояния

$$\frac{dv_1}{dt} = v_2;$$

$$\frac{dv_2}{dt} = v_3;$$

.....

$$\frac{dv_{n-1}}{dt} = v_n.$$

Если $m = 0$, т.е. дифференциальное уравнение имеет вид:

$$A(p)y(t) = b_0 f(t),$$

то n -е уравнение состояния находится просто:

$$\frac{dv_n}{dt} = -\frac{a_0}{a_n}v_1 - \frac{a_1}{a_n}v_2 - \dots - \frac{a_{n-1}}{a_n}v_n + \frac{b_0}{a_n}f.$$

Здесь $v_1 = y$; $v_n = d^{n-1}y/dt^{n-1}$. Матрицы нормальной формы в так называемом каноническом наблюдаемом базисе имеют вид:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -\frac{a_0}{a_n} & -\frac{a_1}{a_n} & -\frac{a_2}{a_n} & \dots & -\frac{a_{n-1}}{a_n} \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{b_0}{a_n} \end{bmatrix}; \quad (2.48)$$

$$C = (1 \ 0 \ 0 \ \dots \ 0)$$

Если оператор при воздействии $B(p)$, т.е. полином числителя B передаточной функции W , имеет степень $0 < m \leq n$, то вид матрицы-столбца изменится; она является решением следующей системы линейных уравнений [21]:

$$LB = G,$$

где B — искомая матрица-столбец (вектор);

$$G = (b_0, b_1, \dots, b_m, 0, \dots, 0)^T;$$

L — треугольная матрица коэффициентов

$$L = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_2 & a_3 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_n & 0 & \dots & 0 \end{bmatrix}.$$

Очевидно, что при таком выборе переменных состояния нет вопроса о вычислении начальных условий, так как $v(0) = y(0)$.

В ряде случаев удобен другой выбор переменных состояния, исключающий необходимость вычисления матрицы B . Матрица состояния A имеет ту же форму (2.48) — так называемую сопровождающую форму Фробениуса (G. Frobenius). Матрица входа B определяется так:

$$B = (0 \ 0 \dots 1)^T,$$

а матрица выхода равна

$$C = (b_0/a_n \ b_1/a_n \dots b_m/a_n \dots 0).$$

В этом каноническом управляемом базисе y уже не является первой по номеру переменной состояния, а выражается как линейная комбинация переменных состояния. Здесь необходим перерасчет начальных условий через обращение матрицы наблюдаемости (см. п. 2.7.1):

$$v(0) = V^{-1}y(0).$$

Коэффициент обхода d отличен от нуля только при равенстве степеней m и n полиномов B и A :

$$d = \begin{cases} b_m/a_n, & m = n; \\ 0, & m < n. \end{cases}$$

До записи матриц нормальной формы при $d \neq 0$ передаточную функцию рекомендуется представлять в виде суммы коэффициента обхода d и правильной дроби:

$$W(s) = d + \frac{B_1(s)}{A(s)},$$

где $m_1 = \deg B_1 < n$, а полином A имеет единичный старший коэффициент.

2.12.2. Построение структурных схем по передаточной функции

Запись дифференциальных уравнений в форме пространства состояний, по существу, является готовым алгоритмом построения графа или схемы реализации динамической системы с использованием n интеграторов. На рис. 2.17 для примера изображен граф, соответствующий системе уравнений (2.18) с матрицами, имеющими вид:

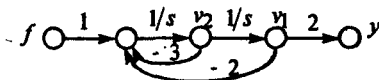


Рис. 2.17. Граф системы второго порядка, построенный по уравнениям в форме пространства состояний

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C = (2 \ 0); \quad D = d = 0$$

Ясно, что в случае канонических базисов, когда матрицы имеют большое число нулей и единиц, реализация упрощается.

Другим способом построения структурных схем по передаточной функции является ее разложение на сумму дробей:

$$W(s) = \frac{B(s)}{A(s)} = \sum_i \frac{C_i}{s - s_i} + K(s),$$

где $C_i = \frac{B(s_i)}{A'(s_i)}$ — коэффициенты разложения (вычеты); s_i — простые (некратные) полюсы передаточной функции; $K(s)$ — полином — неправильная часть исходной передаточной функции.

Суммированию передаточных функций соответствует параллельное соединение звеньев (см. п. 2.10.2). Если все полюсы действительные отрицательные, то слагаемым отвечают передаточные функции типа:

$$\frac{C_i}{s - s_i} = \frac{k_i}{T_i s + 1}.$$

Нулевому полюсу соответствует звено с передаточной функцией

$$\frac{C_i}{s}.$$

В случае комплексных корней $s_{i, \pm 1} = \alpha_i \pm j\omega_i$ двум соседним слагаемым соответствует звено:

$$\frac{C_i}{s - \alpha_i - j\omega_i} + \frac{C_i^*}{s - \alpha_i + j\omega_i} = \frac{k}{T^2 s^2 + 2T\zeta s + 1}; \quad 0 \leq \zeta < 1.$$

Наконец, еще один способ построения типовой структурной схемы по передаточной функции основан на ее факторизованном представлении:

$$W(s) = \frac{B(s)}{A(s)} = K \frac{(s - z_1) \cdots (s - z_m)}{(s - p_1) \cdots (s - p_n)},$$

где z_j ; $j = 1, \dots, m$ — нули; p_i ; $i = 1, \dots, n$ — полюсы передаточной функции. Если все нули и полюсы действительные отрицательные, то получим произведение типовых апериодических и пропорционально-дифференцирующих звеньев (см. п. 2.12.3). В случае комплексных полюсов

сомножителями являются квадратные трехчлены — знаменатели передаточных функций колебательных звеньев. В общем случае имеем

$$W(s) = K \frac{\prod_j (\tau_j s + 1) \prod_\gamma (\tau_\gamma^2 s^2 + 2\tau_\gamma \zeta_\gamma s + 1)}{\prod_i (T_i s + 1) \prod_q (T_q^2 s^2 + 2T_q \zeta_q s + 1)},$$

т. е. любую передаточную функцию можно представить как последовательное соединение типовых звеньев.

2.12.3. Типовые звенья

Как показано выше, любую систему можно представить в виде соединения *типовых динамических звеньев*. Число таких звеньев невелико и определяется типом нулей и полюсов.

Как следует из представления модели в форме пространства состояний, для реализации любой физически осуществимой передаточной функции достаточно двух типов звеньев: *интеграторов* и *усилителей*. Если степень числителя передаточной функции m превышает степень знаменателя n , то необходимо звено *дифференцирующего типа*.

В теории управления состав типовых звеньев несколько расширен исходя из соображения удобства — необходимы звенья, моделирующие часто встречающиеся случаи, а также позволяющие представление передаточных функций общего вида последовательным и параллельным соединением типовых звеньев. Исходя из этого дополнительно рассматривают типовые звенья со следующими передаточными функциями:

- $W(s) = \frac{k}{Ts + 1}$ — *апериодическое звено первого порядка* (k — коэффициент усиления; T — постоянная времени);
- $W(s) = k(Ts + 1)$ — *пропорционально-дифференцирующее звено*;
- $W(s) = \frac{k}{T^2 s^2 + 2T\zeta s + 1}$ — *звено второго порядка* (ζ — коэффициент

демпфирования). При $\zeta \geq 1$ имеем *апериодическое звено второго порядка*; при $0 < \zeta < 1$ — *колебательное звено*; при $\zeta = 0$ — *консервативное звено*. Используют и некоторые другие типы звеньев. В классической теории автоматического регулирования изучению свойств типовых звеньев уделяется большое значение.

2.13. МОДЕЛИ СИСТЕМ УПРАВЛЕНИЯ ВЫСШИХ УРОВНЕЙ

Модели в терминах вход-выход (см. § 2.1) или в форме систем дифференциальных уравнений (см. § 2.5) описывают в свернутой или развернутой формах системы нулевого уровня причинно-следственной интеграции ($L = 0$).

Модели систем первого уровня ($L=1$) образуются как ориентированная взаимосвязь элементов — систем нулевого уровня. На рис. 2.18 изображена топология модели системы управления первого уровня в виде диаграммы графа, в блоках которого вписаны обозначения векторов внутренних переменных x^m и полиномиальных матриц D^m , B^m , C^m ; $m = 1, 2, 3$, систем дифференциальных уравнений. В частном случае блоки описываются уравнениями в форме пространства состояний.

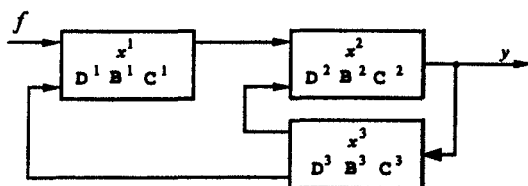


Рис. 2.18. Модель первого уровня в форме блок-графа

Описание моделей высших уровней осуществляется рекурсивно — модель L -го уровня строится как ориентированная взаимосвязь подсистем $(L-1)$ -го уровня (см. п. 1.9.2).

Классические структурные схемы, S -графы и сигнальные графы (см. § 2.8) являются частными случаями моделей первого уровня, построенных на свернутых моделях одномерных звеньев. Для большинства практически важных задач проектирования систем управления эти модели являются наиболее удобными. Для них разработаны правила эквивалентных преобразований (алгебра блок-диаграмм), топологические формулы составления выражений для передаточных функций систем по передаточным функциям звеньев, частотные методы анализа устойчивости и чувствительности, а также методы синтеза.

Методы расчета многомерных систем управления обычно рассматривают частный случай моделей в форме векторно-матричных графов, вершинам которых соответствуют блоки с одним и тем же числом входов и выходов. Блоки таких графов описываются матрицами (A^m , B^m , C^m) уравнений в форме пространства состояний при применении временных методов расчета или передаточными матрицами $W^m(s)$ и комп-

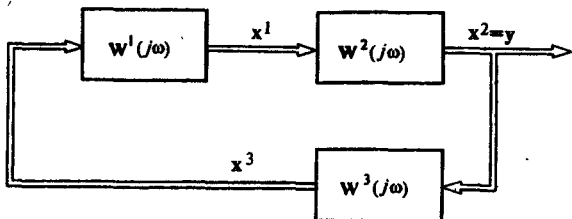


Рис. 2.19. Многомерная система

лексными матрицами частотных характеристик $W^m(j\omega)$ — при применении частотных методов (рис. 2.19).

2.14. МОДЕЛИ СРЕДЫ И РАСШИРЕННОЙ СИСТЕМЫ

Модель среды на входе системы управления M_F — это совокупность воздействий f_r ; $r = 1, \dots, l$. Совокупность модели среды на входе системы M_F и модели системы со связями со средой M_{YSF} образует модель расширенной системы M_{RS} (см. рис. 1.18).

Воздействия могут задаваться как функции времени. Наиболее часто принимаются следующие типы воздействий:

$f_r(t) = \delta(t)$; $F_r(s) = 1$ — единичная δ -функция;

$f_r(t) = 1(t)$; $F_r(s) = 1/s$ — единичное ступенчатое воздействие;

$f_r(t) = \left(\frac{t^\lambda}{\lambda!}\right) 1(t)$; $F_r(s) = \frac{1}{s^{\lambda+1}}$; $\lambda = 1, 2, \dots$ — степенное воздействие

(при $\lambda = 0$ имеем единичное ступенчатое воздействие); (2.49)

$f(t) = (\sin \omega t) 1(t)$; $F(s) = \frac{\omega}{(\omega^2 + s^2)}$ — гармоническое воздействие

единичной амплитуды. (2.50)

Другим способом задания модели среды M_F является подключение ко входу системы так называемых формирующих фильтров. Некоторое «воздействие-первопричина» $f_0(t)$ преобразуется так, чтобы получить требуемое воздействие $f_r(t)$ на входе системы (рис. 2.20).

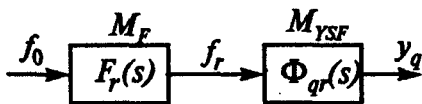


Рис. 2.20. Модель расширенной системы M_{RS}

Если в качестве воздействия $f_0(t)$ принять δ -функцию, то передаточная функция формирующего фильтра равна изображению воздействия:

$$F_r(s) = L\{f_r(t)\}.$$

Если же принять $f_0(t) = 1(t)$, то передаточная функция формирующего фильтра равна

$$F_r(s) = sL\{f_r(t)\}.$$

Пусть, например, на вход системы подано гармоническое воздействие среды (2.50). Если принять за «воздействие-первопричину» $f_0 = 1(t)$, то ко входу системы подключается формирующий фильтр с передаточной функцией

$$F_r(s) = \frac{\omega s}{(\omega^2 + s^2)}.$$

2.15. НЕОПРЕДЕЛЕННОСТЬ МОДЕЛЕЙ СИСТЕМ УПРАВЛЕНИЯ

Математические модели не отражают исчерпывающим образом динамические свойства систем управления в силу идеализаций и упрощений, неизбежных при моделировании, неточной реализации алгоритмов управления и изменений характеристик объектов и других элементов в процессе эксплуатации. Если изменения характеристик происходят достаточно медленно по сравнению с длительностью процессов управления, то вместо нестационарных моделей (например, дифференциальных уравнений с переменными коэффициентами) можно рассматривать множества стационарных моделей.

Модели систем управления строятся для строго оговоренных условий взаимодействия со средой, и их адекватность оригиналам определяется и характеристиками воздействий. Значения параметров, структура операторов и класс операторов зависят от амплитуд изменения и частотного спектра сигналов.

Линейные модели обычно строят для малых отклонений переменных от выбранных установившихся режимов. Если амплитуды сигналов превышают некоторое определенное значение A , то приходится строить нелинейные модели, как правило, учитывающие всевозможные ограничения в реальных элементах. Иногда область адекватности линейных моделей ограничивается малыми амплитудами α , для которых следует учитывать такие нелинейные явления, как зона нечувствительности, сухое трение и др.

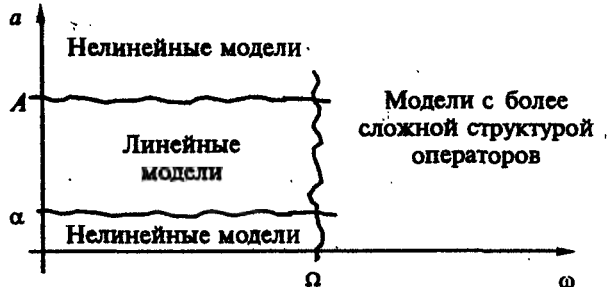


Рис. 2.21. Иллюстрация области адекватности модели

Выбранные структуры операторов (порядки дифференциальных уравнений) обеспечивают адекватность моделей по отношению к сигналам, частоты которых не превышают заданного предела. Границу области адекватности Ω обычно удается несколько расширить путем усложнения структуры операторов.

На рис. 2.21 иллюстрируется область адекватности моделей на плоскости амплитуд a и частот ω сигналов.

Таким образом, модели систем управления оказываются не полностью определенными. Здесь выделяется несколько ситуаций, которые можно систематизировать с помощью введенных ранее рангов неопределенности моделей $R = 0, 1, 2, 3$.

Прежде всего заметим, что менее определенные модели низких рангов R' представляют собой множества более определенных моделей высших рангов R'' :

$$M(R') \subseteq \{M(R'')\}; R' < R''.$$

Модель второго ранга $M(2)$ есть множество моделей третьего ранга $M(3)$, элементы которого различаются значениями параметров. На рис. 2.22, а условно изображен сигнальный граф G , причем, подграф G' отделен полностью. Выделенной дуге (а, б) соответствует передаточная функция, параметры которой принадлежат заданному множеству.

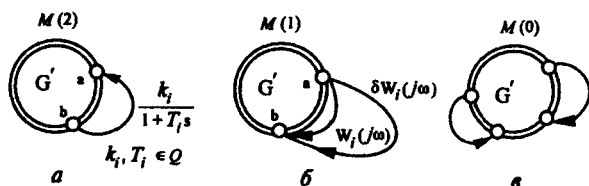


Рис. 2.22. Модели различных рангов неопределенности

ству Q . Следовательно, имеет место множество моделей с различными параметрами — это случай структурированной неопределенности.

Модель первого ранга $M(1)$ является множеством моделей второго ранга $\{M(2)\}$, у которых одна и та же топология, однако структуры операторов могут различаться. В общем случае элементы содержат неструктурированную неопределенность, например, в виде аддитивных вариаций передаточных функций или частотных характеристик $\delta W_i(j\omega)$ дуг графа (рис. 2.22, б).

Модель нулевого ранга $M(0)$ представляет собой множество моделей первого ранга $\{M(1)\}$ с различной топологией. В простейшем случае графы различаются местом включения одной дуги (рис. 2.22, в), например дуги, моделирующей регулятор или корректирующее устройство.

При интерпретации результатов анализа и синтеза необходимо всегда иметь в виду неполную определенность моделей и учитывать ограниченность области их адекватности. Анализ, наряду с выявлением основных свойств поведения систем управления, должен включать и исследование чувствительности характеристик к вариациям параметров, структур операторов и топологии систем. Процедуры синтеза должны обеспечивать робастность систем управления, когда вариации характеристик элементов мало сказываются на процессах управления.

Глава 3 АНАЛИЗ СИСТЕМ УПРАВЛЕНИЯ

3.1. ЗАДАЧИ АНАЛИЗА

Анализ систем управления состоит в изучении их общесистемных свойств, условий выполнения ими своих функций и достижения заданных целей. В результате анализа качественно констатируются свойства поведения систем в целом и количественно оценивается степень удовлетворения требований к процессам управления. Анализ в широком смысле должен также объяснять, почему система с определенной структурой причинно-следственных взаимосвязей элементов с известными характеристиками имеет то или иное поведение.

Безусловными требованиями к свойствам поведения систем управления являются:

- устойчивость движений;
- инвариантность управляемой переменной к возмущениям и ковариантность с задающим воздействием;

- грубость (параметрическая инвариантность, робастность), т.е. ограниченная чувствительность свойств системы к вариациям характеристик элементов.

Основными задачами анализа систем управления, обычно решаемыми методами теории управления, являются:

- установление фактов устойчивости, инвариантности и робастности;
- построение характеристик и вычисление показателей качества;
- вывод об удовлетворительном (или неудовлетворительном) поведении системы.

3.2. АНАЛИЗ УСТОЙЧИВОСТИ

Устойчивость по начальным условиям (по Ляпунову) — свойство, присущее системе. Если система устойчива, то затухают все составляющие свободных движений, вызванных любыми ненулевыми начальными условиями. Свойство устойчивости линейных непрерывных систем анализируется по моделям типа M_s (см. рис. 1.18) в форме однородных дифференциальных уравнений n -го порядка

$$A(p)y(t) \equiv a_n y^{(n)} + \dots + a_1 y' + a_0 y = 0 \quad (3.1)$$

ли систем дифференциальных уравнений в форме пространства состояний

$$\frac{d\mathbf{v}}{dt} = \mathbf{A}\mathbf{v}.$$

3.2.1. Свободные движения

Преобразуя дифференциальное уравнение (3.1) по Лапласу с учетом начальных условий: $y(0), y'(0), \dots, y^{(n-1)}(0)$, получим

$$A(s)Y(s) = A_n(s), \quad (3.2)$$

где $A_n(s)$ — полином, коэффициенты которого зависят от начальных условий. Из алгебраического уравнения (3.2) легко получить изображение решения уравнения (3.1)

$$Y_{\text{св}}(s) = A_n(s)/A(s).$$

В том случае, когда характеристический полином системы $A(s)$ имеет только простые корни s_i ; $i = 1, \dots, n$, выражение для свободных движений имеет вид

$$y_{\text{св}}(t) = \sum_{i=1}^n \frac{A_n(s_i)}{A'(s_i)} e^{s_i t} = \sum_{i=1}^n C_i e^{s_i t}. \quad (3.3)$$

Здесь знак $'$ означает дифференцирование полинома по s . Если корни полинома $A(s)$ кратные, то вместо коэффициентов C_i в выражении (3.3) появятся полиномы от t со степенями ниже кратности корня s_i .

3.2.2. Условие устойчивости

Как следует из выражения (3.3), необходимым и достаточным условием затухания экспонент является отрицательность действительных частей корней

$$\forall i=1, \dots, n: \operatorname{Re} s_i < 0. \quad (3.4)$$

На рис. 3.1 изображен пример расположения корней характеристического полинома асимптотически устойчивой системы пятого порядка на комплексной плоскости. Все корни находятся в открытой левой полуплоскости, т.е. строго левее мнимой оси. Поэтому часто говорят, что для асимптотической устойчивости системы необходимо и достаточно, чтобы все корни характеристического полинома были левыми. Это же условие справедливо и для собственных значений матрицы состояний A .

Заметим, что если характеристический полином имеет простые корни на мнимой оси, то имеет место устойчивость по Ляпунову (но не асимптотическая).

Вынужденные движения неавтономных линейных систем представляются как сумма установившихся движений, определяемых полюсами изображений воздействий и переходных процессов из-за посленулевых начальных условий, вызванных внезапным приложением воздействий. Если системы асимптотически устойчивы, то с течением времени процессы стремятся к установившимся

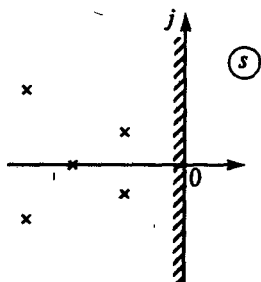


Рис. 3.1. Пример расположения корней устойчивой системы

$$\lim_{t \rightarrow \infty} y(t) = y_{\text{уст}}(t).$$

3.2.3. Устойчивость вход-выход

Система называется устойчивой по входу, если при любом ограниченном воздействии $f(t)$ ее реакция $y(t)$ является ограниченной [57]. Устойчивость по входу характеризует свойство оператора преобразования вход-выход и анализируется по моделям M_{YF} (см. рис.1.18).

Устойчивость вход-выход в случае моделей рассматриваемого класса имеет место, если:

- система устойчива по начальным условиям, т.е. корни ее характеристического полинома находятся в левой полуплоскости;
- передаточная функция системы осуществима (физически реализуема), т. е. степень полинома числителя не превышает степень полинома знаменателя.

3.3. КРИТЕРИИ УСТОЙЧИВОСТИ

Для выявления устойчивости не обязательно знание корней, хотя в условиях применения ЭВМ их вычисление во многих случаях не представляет особых трудностей.

Для суждения об устойчивости системы или звена, не вычисляя корней характеристического полинома, применяют *критерии устойчивости*, которые с помощью относительно простых вычислений позволяют установить, лежат ли все корни в левой полуплоскости.

Различают алгебраические и частотные критерии устойчивости. К алгебраическим относятся критерии Гурвица (А. Hurwitz) и Рауса (Е. Routh), а к частотным — критерии Михайлова и Найквиста (Н. Nyquist).

3.3.1. Необходимое условие устойчивости

При определении устойчивости по характеристическому полиному прежде всего следует проверить выполнение необходимого условия: для того, чтобы все корни полинома имели отрицательные действительные части, все его коэффициенты должны быть одного знака (положительными).

Действительно, представим полином $A(s)$ с единичным старшим коэффициентом как произведение двучленов $(s - s_i)$, если $s_i = \alpha_i$, и трехчленов $[s^2 - 2\alpha_i s + (\alpha_i^2 + \omega_i^2)]$, если $s_i = \alpha_i \pm j\omega_i$. В случае $\alpha_i < 0$; $i = 1, \dots, n$ коэффициенты всех двух- и трехчленов положительны. Следовательно, положительны и все коэффициенты полинома $A(s)$. Для систем первого ($n=1$) и второго ($n=2$) порядка, это условие устойчивости является и достаточным; для $n \geq 3$ необходимо выполнение дополнительных условий.

Типовое аperiodическое звено первого порядка ($n = 1$) устойчиво при $T > 0$; устойчивы и звенья второго порядка при $T > 0$, $\zeta > 0$. Интегрирующее ($n = 1$) и консервативное ($n = 2$) звенья не удовлетворяют условию положительности всех коэффициентов. Они имеют корни на мнимой оси. Это соответствует устойчивости по начальному состоянию (по Ляпунову); однако нет асимптотической устойчивости. Принято говорить, что звенья или системы, имеющие некрратные корни характеристического полинома на мнимой оси (а остальные — левые), находятся на границе устойчивости. Такие системы являются *негрубыми* — они теряют устойчивость при сколь угодно малых изменениях параметров.

3.3.2. Алгебраические критерии

Пусть характеристический полином звена или системы автоматического управления имеет вид:

$$A(s) = a_0 + a_1 s + \dots + a_{n-1} s^{n-1} + a_n s^n.$$

Критерий Гурвица. Для асимптотической устойчивости необходимо и достаточно, чтобы при $a_n > 0$ все диагональные определители матрицы Гурвица были положительны.

Матрица Гурвица составляется из коэффициентов характеристического полинома $A(s)$ следующим образом:

$$H = \begin{bmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \dots & 0 & 0 \\ a_n & a_{n-2} & a_{n-4} & \dots & 0 & 0 \\ 0 & a_{n-1} & a_{n-3} & \dots & 0 & 0 \\ 0 & a_n & a_{n-2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_1 & 0 \\ 0 & 0 & 0 & \dots & a_2 & a_0 \end{bmatrix}.$$

Это квадратная матрица ($n \times n$). Ее диагональными определителями являются

$$\Delta_1 = a_{n-1};$$

$$\Delta_2 = a_{n-1}a_{n-2} - a_n a_{n-3};$$

.....

$$\Delta_n = a_0 \Delta_{n-1}.$$

Если выполнено необходимое условие положительности всех коэффициентов полинома, то вычисляются второй, последующие и предпоследний определители. Очевидно, последний определитель n -го порядка Δ_n положителен при положительности определителя Δ_{n-1} .

Например, для системы третьего порядка

$$A(s) = a_0 + a_1 s + a_2 s^2 + a_3 s^3 \quad (3.4)$$

матрица Гурвица имеет вид

$$H = \begin{pmatrix} a_2 & a_0 & 0 \\ a_3 & a_1 & 0 \\ 0 & a_2 & a_0 \end{pmatrix}.$$

Если выполнено необходимое условие положительности коэффициентов полинома $A(s)$, то следует проверять только знак определителя

$$\Delta_2 = a_1 a_2 - a_0 a_3. \quad (3.5)$$

Для устойчивости системы третьего порядка произведение средних коэффициентов характеристического полинома должно быть больше произведения крайних.

С помощью критерия Гурвица можно строить границы устойчивости и выделять области устойчивости в пространстве коэффициентов полиномов или параметров системы автоматического управления.

Рассмотрим пример системы третьего порядка с характеристическим полиномом (3.4). Пусть все коэффициенты положительны. Проведем замену переменной, не приводящую к изменению расположения корней полинома относительно мнимой оси:

$$s = \Omega \tilde{s}; \quad \Omega = \sqrt[3]{a_0 / a_3},$$

и получим другой полином

$$\tilde{A}(\tilde{s}) = 1 + \beta \tilde{s} + \alpha \tilde{s}^2 + \tilde{s}^3,$$

коэффициенты которого:

$$\alpha = \frac{a_2}{a_0} \Omega^2; \quad \beta = \frac{a_1}{a_0} \Omega$$

называют параметрами Вышнеградского. Границе устойчивости в плоскости параметров α, β соответствует гипербола Вышнеградского $\alpha\beta = 1$ (рис. 3.2). Параметр нормировки Ω называется среднегеометрическим корнем, поскольку $a_0/a_3 = s_1 s_2 s_3 (-1)^3$.



Рис. 3.2. Диаграмма Вышнеградского (фрагмент)

Для систем высоких порядков критерий Гурвица не очень удобен — многократное вычисление определителей становится трудоемким и избыточным. Тогда предпочтительнее применение критерия Рауса, имеющего также алгоритмическую форму. Этот критерий позволяет быстро определить устойчивость системы, если имеется ее характеристический полином $A(s)$ и заданы численно его коэффициенты.

Критерий Рауса наиболее экономичен по объему вычислений в сравнении с другими критериями. Он широко применяется для анализа влияния параметров системы на ее устойчивость с использованием вычислительных машин, поскольку алгоритм вычислений удобен для программирования.

3.3.3. Частотный критерий Михайлова

Критерий Михайлова базируется на *принципе аргумента* [47]. Выражение для характеристического полинома $A(s)$ рассматривается как функция комплексного переменного, принимающего значения на положительной мнимой полуоси. Критерий сводится к анализу изменения аргумента функции $A(j\omega)$.

Согласно критерию Михайлова, для устойчивости системы необходимо и достаточно, чтобы годограф вектора $A(j\omega)$, начинаясь при $\omega = 0$ на действительной положительной полуоси, с ростом ω от нуля до бесконечности обходил последовательно в положительном направлении (против часовой стрелки) n квадрантов, где n — порядок системы:

$$\delta \arg A(j\omega) = n \frac{\pi}{2} \quad 0 \leq \omega < \infty$$

3.4. УСТОЙЧИВОСТЬ СИСТЕМ С ТИПОВОЙ СТРУКТУРОЙ

Рассмотрим системы, образованные последовательным и параллельным соединениями звеньев, а также системы с обратной связью.

Будем иметь в виду асимптотическую устойчивость по начальному состоянию, которая устанавливается по модели собственно системы M_s .

3.4.1. Устойчивость систем без контуров

Как показано в § 2.10, характеристический полином системы в случае последовательного и параллельного соединений звеньев равен произведению характеристических полиномов звеньев

$$A_3(s) = A_1(s)A_2(s).$$

Отсюда следует вывод: необходимым и достаточным условием устойчивости системы, образованной последовательным и параллельным соединениями звеньев, является устойчивость всех звеньев. Вывод сохраняет силу при любом числе звеньев, включенных параллельно и последовательно, т. е. для любых *бесконтурных* графов.

3.4.2. Устойчивость одноконтурных систем

Характеристический полином замкнутой системы, образованной соединением звеньев с обратной связью (см. рис. 2.9, в),

$$A_3(s) = A_1(s)A_2(s) + B_1(s)B_2(s).$$

Для одноконтурных систем, образованных любым числом звеньев с передаточными функциями

$$W_\gamma(s) = B_\gamma(s)/A_\gamma(s); \quad \gamma = 1, 2, \dots,$$

характеристический полином записывается в виде суммы

$$A_3(s) = A_p(s) + B_p(s) \quad (3.6)$$

полиномов знаменателя и числителя передаточной функции разомкнутой системы

$$W_p(s) = \prod_{\gamma} W_\gamma(s).$$

О расположении корней полинома (3.6), полученного суммированием двух полиномов, в общем случае без предварительных вычислений ничего сказать нельзя. Необходимо либо вычислить корни $A_3(s)$, либо применить какой-либо критерий устойчивости. Вместе с тем следует указать на два важных случая.

Если полиномы A_p и B_p имеют нетривиальный общий делитель — полином A_d , т. е. передаточная функция разомкнутой системы W_p имеет диполи, то при замыкании системы соответствующие корни характеристического полинома не перемещаются. Действительно, из выражения

$$A_3 = A_d (A_{p1} + B_{p1})$$

следует, что корни полинома A_d являются и корнями полинома A_3 . Таким образом, необходимое условие устойчивости замкнутой системы — все корни наибольшего общего делителя

$$A_d = \text{НОД}(A_p, B_p)$$

левые. Достаточное условие неустойчивости — наличие у полиномов A_p и B_p общего делителя с правым корнем.

Как показано в § 2.4, наличие общих корней у полиномов числителя и знаменателя передаточной функции соответствует неполной системе. При замыкании такой системы неполная часть своих свойств не изменяет.

Рассмотрим теперь полиномы $A(s)$, $B(s)$ как функции комплексного аргумента s и применим к выражению (3.6) теорему Руше (E. Rouché) [47].

Теорема Руше. Если $A(s)$ и $B(s)$ — две функции, однозначные и аналитические в точках замкнутой кривой C без самопересечений и внутри нее, и если в точках этой кривой выполнено условие

$$|A(s)| > |B(s)|,$$

то внутри C сумма

$$A(s) + B(s)$$

имеет столько же нулей, сколько их имеет функция $A(s)$.

Выберем замкнутую кривую C , образованную мнимой осью ($s = j\omega$) и полуокружностью бесконечного радиуса (рис. 3.3). Так называемый контур Найквиста на комплексной плоскости охватывает правую полуплоскость.

Полиномы A_p и B_p в выражении (3.6) удовлетворяют условиям теоремы Руше об однозначности и непрерывности. Если выполняется условие

$$\forall s \in C: |A_p(s)| > |B_p(s)| \quad (3.7)$$

иначе

$$\forall \omega |A_p(j\omega)| > |B_p(j\omega)|, \quad (3.8)$$

то характеристический полином замкнутой системы A_z имеет в правой полуплоскости столько же корней, что и характеристический полином разомкнутой системы A_p .

Пусть звенья, образующие контур обратной связи, устойчивы, т. е. полиномы A_v имеют только левые корни. Тогда все корни полинома A_p — левые, т. е. находятся вне области, охваченной контуром C . Поэтому при выполнении достаточного условия (3.7) устойчива и замкнутая система.

В силу равенства $|A(j\omega)| = |A(-j\omega)|$ можно ограничиться проверкой условия (3.8) для $s = j\omega$; $\omega \geq 0$.

Как следует из условия (3.8), для устойчивости замкнутого контура отрицательной обратной связи, образованного из устойчивых звеньев, достаточно

$$|W_p(j\omega)| < 1, \quad \omega \geq 0. \quad (3.9)$$

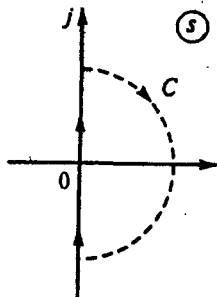


Рис. 3.3. Контур Найквиста

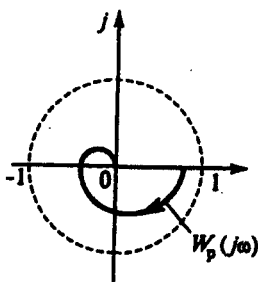


Рис. 3.4. Иллюстрация достаточного условия устойчивости

Для устойчивости замкнутой системы, устойчивой в разомкнутом состоянии, достаточно, чтобы амплитудно-фазовая характеристика разомкнутой системы принадлежала единичному кругу (рис.3.4) с центром в начале координат.

Если звенья образуют контур положительной обратной связи, формула для характеристического полинома замкнутой системы имеет вид

$$A_3 = A_p - B_p.$$

И в этом случае условие (3.9) гарантирует сохранение числа корней в правой полуплоскости, т.е. даже смена знака обратной связи не нарушает устойчивости (или неустойчивости). Условия (3.7) — (3.9) требуют малого усиления контура на всех частотах. Если усиление контура мало, его замыкание или размыкание хотя и приводит к некоторому перемещению корней, но не так сильно, чтобы они переходили через мнимую ось (контур C на комплексной плоскости).

Это условие не является необходимым для устойчивости — на некоторых частотах усиление может быть большим. Именно большое усиление контура (много больше единицы) на частотах возмущений позволяет добиться ослабления их влияния на управляемую переменную, т.е. инвариантности управляемой переменной замкнутой устойчивой системы к непосредственно неизмеряемым возмущениям.

3.4.3. Приближенная зависимость корней характеристического полинома одноконтурной системы от нулей и полюсов передаточной функции разомкнутой системы

При замыкании контура не все корни характеристического полинома перемещаются в одинаковой степени. Это зависит от усиления контура, т.е. значения амплитудно-частотной характеристики $|W_p(j\omega)|$ на

частотах, равных модулям корней. Усиление контура на конкретной частоте зависит как от коэффициента передачи, так и от взаимного расположения нулей и полюсов. Вместе с тем можно указать на некоторые простые приближенные соотношения между корнями характеристического полинома замкнутой системы $A_3(s)$ и нулями и полюсами передаточной функции $W_p(s)$, вытекающие из рассмотрения ЛАЧХ $L_p(\omega)$ [43].

Пусть на множестве частот $\omega \in \Omega_1$ значения амплитудно-частотной характеристики разомкнутого контура много больше единицы

$$\forall \omega \in \Omega_1 : |W_p(j\omega)| \gg 1.$$

Если исключить окрестности нулей $\{z_j\}$ передаточной функции $W_p(s)$, это условие можно распространить на множество значений комплексного аргумента

$$\forall s: |s| \in \Omega_1 : |W_p(s)| \gg 1.$$

Зафиксируем один из простых нулей z_j и сформируем функцию

$$W_{p1}(s) = W_p(s) / (s - z_j),$$

которая не имеет нуля z_j . Для этой функции справедливо

$$\forall s: |s| \in \Omega_1 : |B_p(s)/A_p(s)(s - z_j)| \gg 1$$

и для окрестности выделенного нуля z_j .

Представим характеристическое уравнение замкнутой системы в следующем виде:

$$A_3(s) = \frac{B_p(s)}{s - z_j} \left(\frac{A_p(s)(s - z_j)}{B_p(s)} + (s - z_j) \right) = 0.$$

Поскольку $B_p(s)/(s - z_j) \neq 0$, а первое слагаемое в скобках мало для $|s| \in \Omega_1$, то характеристический полином приближенно равен

$$A_3(s) \approx (s - z_j),$$

т. е. имеет корень z_j .

Отсюда следует, что корни характеристического полинома замкнутой системы приближенно равны нулям передаточной функции разомкнутой системы, модули которых принадлежат множеству частот, где амплитудно-частотная характеристика разомкнутого контура много больше единицы.

Пусть на множестве частот Ω_2 значения амплитудно-частотной характеристики много меньше единицы:

$$\forall \omega \in \Omega_2 : |W_p(j\omega)| \ll 1.$$

Если исключить окрестности полюсов $\{p_i\}$ передаточной функции $W_p(s)$, это условие можно распространить на множество значений комплексного аргумента

$$\forall s: |s| \in \Omega_2 : |W_p(s)| \ll 1.$$

Зафиксируем один из простых полюсов p_i и сформируем функцию

$$W_{p_2}(s) = W_p(s)(s - p_i),$$

которая не имеет полюса p_i . Для этой функции справедливо

$$\forall s \in \Omega_2 : |B_p(s)(s - p_i) / A_p(s)| \ll 1$$

и для окрестности выделенного полюса.

Представим характеристическое уравнение замкнутой системы в следующем виде:

$$A_3(s) = \frac{A_p(s)}{s - p_i} \left((s - p_i) + \frac{B_p(s)(s - p_i)}{A_p(s)} \right) = 0.$$

Очевидно, для $|s| \in \Omega_2$ характеристическое уравнение запишется так:

$$A_3(s) \approx s - p_i = 0.$$

Отсюда следует, что корни характеристического полинома замкнутой системы приближенно равны полюсам передаточной функции разомкнутой системы, модули которых принадлежат множеству частот, где амплитудно-частотная характеристика разомкнутого контура много меньше единицы. Это и понятно: на этих частотах контур практически разомкнут, поэтому соответствующие корни характеристических полиномов замкнутой и разомкнутой систем близки.

Взаимосвязь корней характеристического полинома замкнутой системы с коэффициентом передачи, полюсами и нулями передаточной функции разомкнутого контура особенно удобно исследовать по асимптотическим ЛАЧХ. Точкам сопряжения асимптот соответствуют модули нулей и полюсов передаточной функции разомкнутой системы.

Практически $|W_p(j\omega)| \gg 1$, если $L_p(\omega) > 16$ дБ, а $|W_p(j\omega)| \ll 1$, если $L_p(\omega) < -16$ дБ.

Пример 1. Рассмотрим систему с отрицательной обратной связью, передаточная функция разомкнутого контура которой

$$W_p(s) = \frac{30(s+1)}{s(2s+1)(0,001s+1)(0,005s+1)}.$$

Она имеет нуль $z_1 = -1$ и четыре полюса: $p_1 = 0$; $p_2 = -0,5$; $p_3 = -100$ и $p_4 = -200$. Следовательно, на частоте $\omega_1 = 1 \text{ с}^{-1}$ наклон асимптотической ЛАЧХ меняется на $+20$ дБ/дек, а на частотах $\omega_2 = 0,5 \text{ с}^{-1}$, $\omega_3 = 10 \text{ с}^{-1}$ и $\omega_4 = 200 \text{ с}^{-1}$ — на -20 дБ/дек.

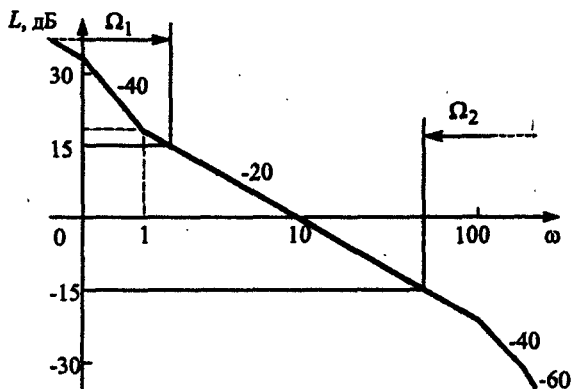


Рис. 3.5. Пример асимптотической ЛАЧХ

На рис. 3.5 изображена асимптотическая ЛАЧХ $L_p(\omega)$. По ней можно заключить, что характеристический полином замкнутой системы $A_3(s)$ будет иметь корень $s_1 \approx -1$, так как на частоте $\omega_1 = 1 \text{ с}^{-1}$ усиление контура больше 16 дБ. Кроме того, $A_3(s)$ будет иметь корни $s_2 \approx -100$ и $s_3 \approx -200$, так как на частотах $\omega > 100 \text{ с}^{-1}$ усиление контура меньше -16 дБ.

Как показывают расчеты, точные корни характеристического полинома замкнутой системы равны: $s_1 = -1,0365$; $s_4 = -20,4953$; $s_2 = -66,45$; $s_3 = -212,5$.

3.5. КРИТЕРИЙ НАЙКВИСТА

Критерий Найквиста является необходимым и достаточным условием устойчивости систем с обратной связью.

Рассмотрим одноконтурную систему с передаточной функцией разомкнутого контура $W_p(s)$. Рациональная функция комплексного аргумента

$$\Delta(s) = 1 + W_p(s)$$

называется *возвратной разностью*. Отметим также, что эта функция — определитель одноконтурного графа с отрицательной обратной связью (см. п. 2.11.2). Возвратная разность равна отношению характеристических полиномов замкнутой и разомкнутой систем

$$\Delta(s) = 1 + \frac{B_p(s)}{A_p(s)} = \frac{A_p(s) + B_p(s)}{A_p(s)} = \frac{A_3(s)}{A_p(s)}.$$

Критерий Найквиста базируется на *принципе аргумента* [47]. Пусть C — произвольный замкнутый контур без самопересечений на плоскости s , а $\Delta(s)$ — рациональная функция s , не имеющая на контуре C ни

нулей, ни полюсов. Разность $z - p$ между количеством нулей и полюсов однозначной функции $\Delta(s)$, заключенных внутри замкнутой кривой C , равна числу полных оборотов, которые делает вокруг начала координат вектор $\Delta(s)$, в то время как точка s описывает контур C по часовой стрелке. При исследовании асимптотической устойчивости в качестве контура C выбирается мнимая ось и полуокружность бесконечного радиуса (контур Найквиста охватывает правую полуплоскость).

Нулями $\Delta(s)$ являются корни характеристического полинома замкнутой системы, а полюсами — корни характеристического полинома разомкнутой системы. Если $p = 0$, т.е. разомкнутая система устойчива, то для устойчивости замкнутой системы ($z = 0$) необходимо и достаточно, чтобы кривая $\Delta(s)$ при движении s вдоль C по часовой стрелке не охватывала начала координат.

Вместо возвратной разности $\Delta(s)$ можно рассматривать *возвратное отношение* — передаточную функцию разомкнутой системы

$$W_p(s) = \Delta(s) - 1.$$

При этом для устойчивости замкнутой системы необходимо и достаточно, чтобы годограф $W_p(s)$ — амплитудно-фазовая характеристика (АФХ) разомкнутой системы — при движении s вдоль C по часовой стрелке не охватывал точку $(-1, j0)$.

На рис. 3.6, а изображен случай устойчивой замкнутой системы, которая устойчива и в разомкнутом состоянии.

Если АФХ разомкнутой системы проходит через критическую точку $(-1, j0)$ на частоте ω_0 , то пара корней характеристического полинома замкнутой системы окажется чисто мнимой $\pm j\omega_0$. Этот случай называют колебательной границей устойчивости (рис. 3.6, б).

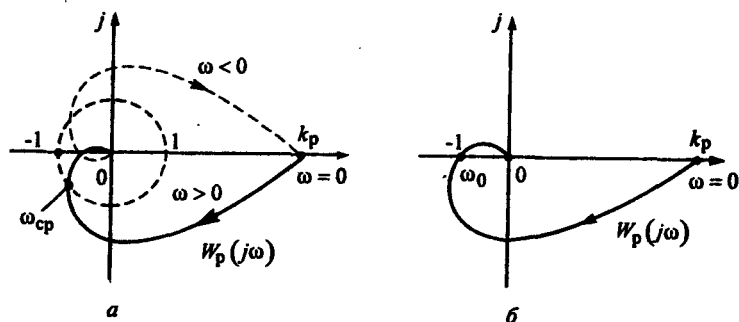


Рис. 3.6. Иллюстрация критерия Найквиста

Рациональные функции $\Delta(s)$ и $W_p(s)$ имеют одни и те же полюсы. Если среди них имеется p правых полюсов, т. е. разомкнутая система является неустойчивой, для устойчивости замкнутой системы ($z = 0$) АФХ $W_p(j\omega)$ должна p раз охватывать точку $(-1, j0)$ против часовой стрелки. В силу симметричности характеристик:

$$\operatorname{Re} W_p(j\omega) = \operatorname{Re} W_p(-j\omega); \quad \operatorname{Im} W_p(j\omega) = -\operatorname{Im} W_p(-j\omega)$$

можно ограничиться рассмотрением $s = j\omega$; $\omega \geq 0$, т. е. половины контура C на комплексной плоскости. Соответственно изменится и формулировка критерия Найквиста.

Для устойчивости замкнутой системы необходимо и достаточно, чтобы АФХ $W_p(j\omega)$ охватывала $p/2$ раз точку $(-1, j0)$ против часовой стрелки, где p — число правых корней характеристического полинома разомкнутой системы (правых полюсов передаточной функции разомкнутой системы).

Очевидно, выполнение достаточного условия устойчивости (3.9) гарантирует, что АФХ (см. рис. 3.4) не охватывает точку $(-1, j0)$. Вместе с тем ясно, что необходимое и достаточное условие Найквиста оставляет большую свободу для формирования $W_p(j\omega)$ при условии устойчивости замкнутой системы.

Пусть для примера передаточная функция разомкнутой системы имеет вид

$$W_p(s) = \frac{k}{Ts - 1},$$

т. е. имеет один правый полюс. Разомкнутая система неустойчива. Для устойчивости замкнутой системы АФХ $W_p(j\omega)$ при изменении ω от нуля до бесконечности должна $1/2$ раза охватить точку $(-1, j0)$ против часовой стрелки. Как видно из рис. 3.7, это возможно при $k > 1$ (замкнутая система устойчива при $k = k_1$ и неустойчива при $k = k_2$). Для перемещения корня из правой полуплоскости в левую необходимо достаточно большое усиление контура.

Передаточную функцию разомкнутой системы во многих случаях удобно представлять в следующем виде:

$$W_p(s) = k_p W_{p0}(s) = k_p \frac{B_{p0}(s)}{A_{p0}(s)},$$

где k_p — коэффициент передачи разомкнутой системы. Он равен отношению младших, отличных от нуля, коэффициентов полиномов числителя и знаменателя. Передаточную функцию W_{p0} назовем нормированной. В частном случае ненулевых нулей и полюсов

$$k_p = W_p(0)$$

является коэффициентом усиления, который характеризует свойства контура по постоянному сигналу.

Рассмотрим рациональную функцию

$$\Delta(s) = \frac{1}{k_p} + W_{p0}(s) = \frac{A_{p0}(s) + k_p B_{p0}(s)}{k_p A_{p0}(s)},$$

числитель которой — характеристический полином замкнутой системы. Для устойчивости замкнутой системы нормированная АФХ $W_{p0}(j\omega)$; $\omega \geq 0$, должна $p/2$ раз охватывать точку $(-1/k_p, j0)$ против часовой стрелки.

Такая формулировка критерия Найквиста упрощает исследование зависимости устойчивости замкнутой системы от коэффициента передачи контура. При изменении k_p нормированная АФХ не изменяется, а критическая точка $(-1/k_p, j0)$ превращается в критический отрезок (луч), как это показано на рис. 3.8. Здесь легко найти критический коэффициент усиления — он соответствует точке пересечения АФХ $W_{p0}(j\omega)$ с критическим отрезком.

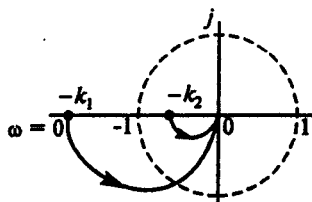


Рис. 3.7. Пример применения критерия Найквиста

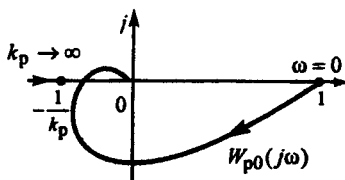
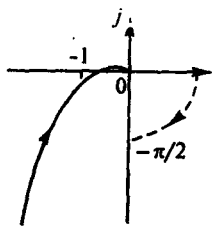


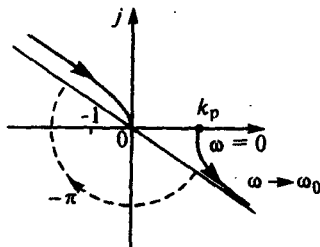
Рис. 3.8. Применение критерия Найквиста для нормированных характеристик

Если передаточная функция разомкнутой системы $W_p(s)$ имеет полюсы на мнимой оси (остальные левые) — нулевой полюс в случае интегратора в составе звеньев или пару мнимых полюсов консервативного звена, то выбор контура C имеет свою специфику. Чтобы принять число корней p разомкнутой системы внутри контура C равным нулю и сохранить формулировку критерия, этот контур обходит полюсы $W_p(s)$ на мнимой оси по полуокружностям бесконечно малого радиуса.

АФХ $W_p(j\omega)$ при значениях, близких к полюсам на мнимой оси, а именно, при их обходе против часовой стрелки по дугам окружности бесконечно малого радиуса, принимает по модулю бесконечно большое значение; аргумент АФХ изменяется на $-\pi$ (по часовой стрелке). Если $\omega \geq 0$, то в случае нулевого полюса аргумент $W_p(j\omega)$ изменяется



a



б

Рис. 3.9. Примеры АФХ для критических случаев

при $\omega = 0$ на $-\pi/2$. На рис. 3.9, *a* в качестве примера изображена АФХ для передаточной функции:

$$W_p(s) = \frac{k_p}{s(T_1s + 1)(T_2s + 1)},$$

а на рис. 3.9, *б* — для

$$W_p(s) = \frac{B_p(s)}{(s^2 + \omega_0^2)A_{p1}(s)}.$$

В случае систем высокого порядка АФХ может иметь сложный вид, затрудняющий подсчет числа охватов критической точки. Для упрощения рекомендуется считать число переходов АФХ через луч $(-\infty, -1)$. Переход снизу вверх считается отрицательным, а сверху вниз — положительным. Для устойчивости замкнутой системы необходимо и достаточно, чтобы сумма переходов через луч равнялась $+p/2$, где p — число правых полюсов передаточной функции разомкнутой системы. Особенно удобно применение критерия Найквиста, а также выявление влияния свойств звеньев на устойчивость, если строятся логарифмические частотные характеристики:

$$L_p(\omega) = 20 \lg |W_p(j\omega)|; \quad \varphi_p(\omega) = \arg W_p(j\omega).$$

Если передаточная функция разомкнутой системы имеет p правых полюсов, для устойчивости замкнутой системы необходимо и достаточно, чтобы в интервале частот, где $L_p > 0$ число переходов $\varphi_p(\omega)$ через линию $-\pi$ равнялось $p/2$.

Положительным считается переход снизу вверх, а отрицательным — сверху вниз.

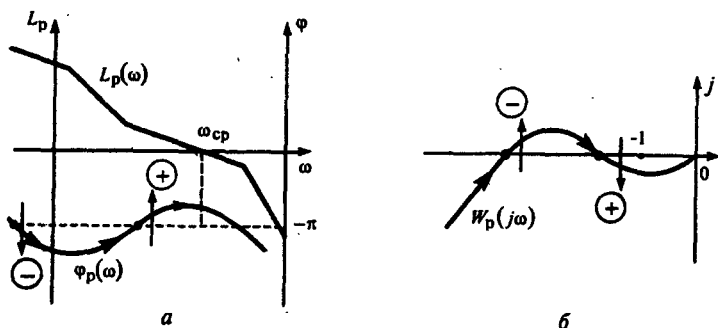


Рис. 3.10. Характеристики условно-устойчивой системы

На рис. 3.10, а изображены логарифмические частотные характеристики, а на рис. 3.10, б — АФХ так называемой условно-устойчивой системы.

Критерий Найквиста можно применить только в том случае, когда выполняется необходимое условие устойчивости — неполная часть системы устойчива (диполи передаточной функции $W_p(s)$ — левые).

Критерий Найквиста физичен. Хорошо видна роль амплитудных и фазовых преобразований, вносимых контуром, на устойчивость системы в целом. Изначальный смысл применения критерия Найквиста заключается не столько в констатации устойчивости, сколько в выявлении роли контура в перемещении корней характеристического полинома системы. На базе этого критерия можно судить о влиянии свойств элементов на характер свободных движений системы в целом.

Практически важное свойство критерия Найквиста заключается также в том, что по нему можно исследовать устойчивость систем с обратными связями на основе экспериментально снятых частотных характеристик звеньев, образующих контур.

3.6. ИНВАРИАНТНОСТЬ СИСТЕМ УПРАВЛЕНИЯ

Под *инвариантностью* систем управления понимается независимость некоторых ее переменных от воздействий. Инвариантность — свойство операторов преобразований; оно анализируется по моделям типа $M_{\text{ИСФ}}$ (см. рис. 1.18).

Теория инвариантности систем управления развивалась в трудах Г. В. Щипанова [92], Н. Н. Лузина [45], Б. Н. Петрова, В. С. Кулебакина и др. [72].

Управляемая переменная системы $y(t)$ в общем случае зависит как от задающего $g(t)$, так и возмущающего $f(t)$ воздействий. При этом переменная выхода $y(t)$ должна быть инвариантной к возмущению и ковариантной с заданием.

3.6.1. Формы инвариантности

Изображение переменной выхода системы при нулевых начальных условиях равно:

$$Y(s) = W_g(s)G(s) + W_f(s)F(s).$$

Основным в теории инвариантности является случай, когда о воздействиях нет никакой информации (в том числе статистического характера). Это значит, что для *абсолютной инвариантности* должно выполняться условие равенства нулю передаточной функции для всех значений комплексного аргумента:

$$\forall s: W_f(s) = 0. \quad (3.10)$$

Абсолютная ковариантность управляемой переменной с задающим воздействием достигается при условии:

$$\forall s: W_g(s) = 1.$$

При выполнении условия (3.10), как это видно из разложения (2.13), оказываются нулевыми установившаяся $y_{уст}(t)$ и переходная $y_{пер}(t)$ составляющие вынужденного движения системы. Однако в системе с ненулевыми предначальными условиями имеют место свободные движения.

Условие абсолютной инвариантности (3.10) — тождественное равенство нулю передаточной функции системы — реализуется редко. Обычно добиваются условий *инвариантности до ϵ*

$$\forall s: |W_f(s)| < \epsilon, \quad (3.11)$$

т. е. ограничивают степень и высоту полинома числителя передаточной функции (операторного полинома при воздействии) [45].

При выполнении условия (3.11) оказываются ограниченными вынужденные составляющие движений.

Близким к инвариантности до ϵ является понятие устойчивости вход-выход (см. п. 3.2.3). Оба понятия связаны со свойством оператора преобразования входных переменных.

Форма инвариантности, когда обеспечивается малость установившейся реакции системы на некоторые типы воздействий, называется

селективной инвариантностью. Селективная инвариантность — свойство расширенной системы; она анализируется по модели типа M_{RS} (см. рис. 1.18).

Условие *селективной абсолютной инвариантности* записывается как условие равенства нулю передаточной функции системы на полюсах s_k изображения возмущения:

$$\forall s_k : W_f(s_k) = 0.$$

Селективная инвариантность до ϵ — ограниченность установившейся реакции системы на воздействие определенного типа — достигается при условии

$$\forall s_k : |W_f(s_k)| < \epsilon.$$

Очевидно, следует стремиться минимизировать значения модуля передаточной функции системы по возмущению в той области комплексной плоскости, где размещаются полюсы изображения воздействия. Как видно из выражения (2.13), малость значений передаточной функции на полюсах возмущения

$$W(s_k) = \frac{B(s_k)}{A(s_k)}; \quad k = 1, \dots, n_F,$$

т. е. сомножителей коэффициентов разложения, означает малость установившихся составляющих реакции.

Существенным условием инвариантности является отсутствие резонанса в системе. Обобщенный резонанс означает совпадение полюсов изображения воздействия с полюсами передаточной функции системы (комплексных частот воздействия с собственными комплексными частотами): $s_k = s_r$.

Для лучшего подавления воздействия необходимо отдалять полюсы передаточной функции от полюсов воздействия.

Ограничение модуля передаточной функции на частотах воздействия приводит к ослаблению вынужденных установившихся реакций. Однако при этом переходная составляющая вынужденных движений, обусловленная послена начальными условиями, может и не ослабляться. Как следует из выражения (2.13), эта составляющая реакции системы будет тем меньше, чем дальше расположены полюсы передаточной функции s_i от полюсов изображения воздействия (это увеличивает $A_F(s_i)$) и чем ближе нули передаточной функции к ее полюсам (это уменьшает $B(s_i)$).

Реакция системы $y(t)$ на задающее воздействие должна быть как можно ближе к $g(t)$. Говорят, что выход системы должен быть ковариантным с заданием. Как следует из выражения (2.13), система хорошо воспроизводит задающее воздействие, если передаточная функция $W_g(s)$ имеет близкие значения на полюсах воздействия:

$$\frac{B_g(s_1)}{A(s_1)} \approx \frac{B_g(s_2)}{A(s_2)} \approx \dots \approx \frac{B_g(s_{n_F})}{A(s_{n_F})}.$$

Это значит, что амплитудно-частотная характеристика системы по каналу воспроизведения должна быть постоянной на спектре задающего сигнала. На рис. 3.11, а изображены графики амплитудно-частотных характеристик абсолютно инвариантной (прямая 1, совпадающая с осью абсцисс) и инвариантной до ϵ (кривая 2) систем. Амплитудно-частотные характеристики абсолютно ковариантной (прямая 1) и ковариантной до ϵ (кривая 2) систем приведены на рис. 3.11, б, где ω_p — резонансная частота; M — показатель колебательности.

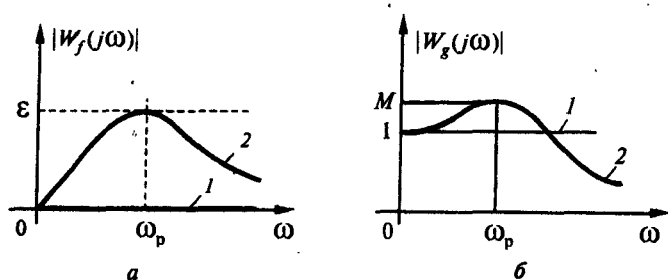


Рис. 3.11. Амплитудно-частотные характеристики систем:
а — инвариантных; б — ковариантных

Все изложенное об инвариантности предполагает, что система асимптотически устойчива, т. е. свободная составляющая движений $y_{св}(t)$ и вынужденная переходная составляющая $y_{пер}(t)$ со временем затухают до нуля.

3.6.2. Селективная инвариантность к степенным воздействиям

Наиболее часто обеспечивается селективная инвариантность к постоянному воздействию. Изображение постоянного воздействия

$$F(s) = f_0 / s$$

имеет нулевой полюс. Установившаяся вынужденная составляющая $y_{уст}(t)$ реакции системы определяется по формуле (2.13):

$$y_{уст}(t) = \frac{B(0)}{A(0)} f_0 = W_f(0) f_0.$$

Чем меньше статический коэффициент усиления системы по каналу возмущающего воздействия, тем меньше установившееся отклонение управляемой переменной — достигается селективная инвариантность до ϵ . Если передаточная функция имеет нулевой нуль

$$W_f(0) = 0,$$

то имеет место нулевая установившаяся реакция на постоянное воздействие — селективная абсолютная инвариантность. В этом случае также говорят, что система *астатическая*.

Если $f(t) = f_1 t$, т. е. изображение

$$F(s) = f_1 / s^2$$

имеет двукратный нулевой полюс, то вместо (2.13) необходимо пользоваться более общим выражением для вынужденной установившейся составляющей $y_{уст}(t)$. Однако нас интересует предельное значение реакции системы, которое легко найти по теореме о конечном значении оригинала

$$\lim_{t \rightarrow \infty} y(t) = \lim_{s \rightarrow 0} s Y(s).$$

Для рассматриваемого случая имеем

$$y_{уст} = \lim_{s \rightarrow 0} s W_f(s) f_1 / s^2.$$

Если передаточная функция системы имеет нулевой нуль

$$W_f(s) = \frac{s B_1(s)}{A(s)},$$

то установившееся отклонение выхода системы из-за возмущения, изменяющегося с постоянной скоростью, равно постоянной величине

$$y_{уст} = \frac{B_1(0)}{A(0)} f_1.$$

Для селективной абсолютной инвариантности передаточная функция $W_f(s)$ должна иметь двукратный нулевой нуль; говорят об астатизме второго порядка.

Рассмотрим общий случай степенного воздействия (2.49), изображение которого имеет $(\lambda + 1)$ -кратный нулевой полюс. Пусть передаточная функция системы по возмущению имеет ν -кратный нулевой нуль

$$W_f(s) = \frac{s^\nu B_1(s)}{A(s)}.$$

Предельное значение реакции системы на степенное воздействие

$$y_{уст} = \lim_{s \rightarrow 0} s \frac{s^\nu B_1(s)}{A(s)} \cdot \frac{f_\lambda}{s^{\lambda+1}} = \frac{B_1(0)}{A(0)} f_\lambda \lim_{s \rightarrow 0} s^{\nu-\lambda}.$$

Отсюда следует: при $\nu > \lambda$ — в системе нулевая установившаяся реакция; при $\nu = \lambda$ — установившаяся реакция постоянна; при $\nu < \lambda$ — установившаяся реакция на возмущение имеет бесконечное значение (система управления неработоспособна).

3.6.3. Селективная инвариантность к гармоническому воздействию

Пусть к системе управления приложено гармоническое возмущение

$$f(t) = a_f \sin \omega_f t,$$

изображение которого

$$F(s) = \frac{a_f \omega_f}{s^2 + \omega_f^2}$$

имеет чисто мнимые полюсы $s_{1,2} = \pm j\omega_f$. Установившаяся реакция линейной системы на гармоническое воздействие является гармонической с той же частотой ω_f , но с другой амплитудой и фазой

$$y_{уст}(t) = a_y \sin(\omega_f t + \varphi_y),$$

где

$$a_y = |W_f(j\omega_f)| a_f;$$

$$\varphi_y = \arg W_f(j\omega_f).$$

Если модуль комплексной частотной характеристики системы управления по каналу возмущения мал, то имеет место селективная инвариантность до ϵ . Для селективной абсолютной инвариантности переменной выхода полином числителя $B(s)$ передаточной функции $W_f(s)$ должен иметь сомножитель $(s^2 + \omega_f^2)$, т. е. пару мнимых корней $\pm j\omega_f$.

3.7. ИНВАРИАНТНОСТЬ СИСТЕМ С ТИПОВОЙ СТРУКТУРОЙ

Рассмотрим, как может быть обеспечена инвариантность (ковариантность) в системах управления, имеющих различные структуры, т. е. реализующих различные принципы управления (см. §1.4).

3.7.1. Последовательная компенсация воздействия

Для подавления возмущения, приложенного к объекту, теоретически можно применить последовательную компенсацию (рис. 3.12). Абсолютная инвариантность переменной выхода объекта имеет место, если выполняется условие

$$W_{of}(s)W_k(s) \equiv 0,$$

где W_{of} , W_k — передаточные функции объекта по каналу возмущения и компенсатора. Отсюда получим тривиальное условие нулевого усиления компенсатора на всех частотах. Инвариантность до ε достигается, если

$$\forall s: |W_{of}(s)W_k(s)| < \varepsilon.$$

Селективная абсолютная инвариантность к воздействию с полюсами изображения s_k имеет место при

$$\forall s_k: W_{of}(s_k)W_k(s_k) = 0.$$

Это означает, что передаточная функция компенсатора (или объекта) должна иметь нули, равные полюсам возмущения. Например, в случае постоянного возмущения

$$F(s) = f_0 / s$$

условие селективной абсолютной инвариантности сводится к наличию нулевого нуля у передаточной функции $W_k(s)$. Это обеспечит нулевую установившуюся реакцию объекта при любом уровне f_0 постоянного

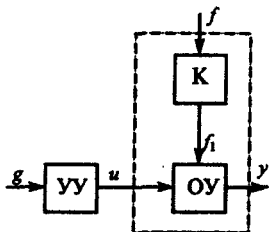


Рис. 3.12. Последовательная компенсация возмущения

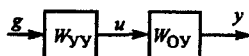


Рис. 3.13. Обеспечение ковариантности

воздействия. Если сам объект по каналу возмущения имеет интегрирующие свойства, т.е. передаточная функция $W_{\text{ж}}(s)$ имеет нулевой полюс, то передаточная функция $W_{\text{к}}(s)$ должна иметь двукратный нулевой нуль.

Следует иметь в виду, что на выходе объекта будут наблюдаться переходные процессы из-за посленаачальных условий, вызванных ступенчатым изменением воздействия. Другими причинами изменения выхода объекта могут быть ненулевые преднаачальные условия и другие, некомпенсированные воздействия.

Хотя последовательная компенсация возмущения выглядит весьма привлекательно, ее реализация обычно связана с изменением конструкции самого объекта, технологии управляемого процесса. В качестве простого примера ослабления действия помех, приложенных к объектам электрической природы, можно указать на экранирование проводов от внешних магнитных полей.

Для достижения ковариантности управляемой переменной $y(t)$ с задающим воздействием $g(t)$ управляющее устройство соединяют последовательно с объектом (рис. 3.13). Выполнение условий абсолютной ковариантности:

$$W_{\text{ou}}(s)W_{\text{yy}}(s) \equiv 1; \quad (3.12)$$

$$W_{\text{yy}}(s) \equiv 1/W_{\text{ou}}(s) \quad (3.13)$$

наталкивается на проблему реализуемости управляющего устройства. Так как объект обычно инерционный, полином знаменателя его передаточной функции по каналу управления $W_{\text{ou}}(s)$ имеет степень выше, чем степень полинома числителя. Следовательно, для реализации условия (3.12) передаточная функция (3.13) должна иметь более высокую степень числителя, чем степень знаменателя.

Селективная абсолютная ковариантность вполне реализуема, так как условие (3.12) в этом случае должно выполняться только на полюсах задающего воздействия, т.е. при $s = s_k$, $k = 1, 2, \dots, n_G$. Например, если известно, что задание имеет постоянные уровни, то при условии

$$W_{\text{ou}}(0)W_{\text{yy}}(0) = 1$$

установившиеся значения переменной выхода объекта будут в точности равны заданию. Необходимо, чтобы коэффициент усиления управляющего устройства был равен величине, обратно пропорциональной коэффициенту усиления объекта по каналу управления.

3.7.2. Параллельная компенсация воздействия

Параллельная компенсация (рис. 3.14) является мощным средством обеспечения инвариантности к основному возмущению на объект, если это возмущение может быть измерено.

Абсолютная инвариантность выхода объекта к возмущению достигается при условии

$$W_{of}(s) + W_{ou}(s)W_{uf}(s) \equiv 0$$

или

$$W_{uf}(s) = -W_{of}(s)/W_{ou}(s).$$

Если раскрыть передаточные функции объекта до полиномов

$$W_{of}(s) = \frac{B_{of}(s)}{A_o(s)}; \quad W_{ou}(s) = \frac{B_{ou}(s)}{A_o(s)},$$

то получим передаточную функцию управляющего устройства по каналу возмущения

$$W_{uf}(s) = -\frac{B_{of}(s)}{B_{ou}(s)}.$$

Здесь принято, что знаменатели W_{of} и W_{ou} совпадают. Поскольку во многих случаях инерционные свойства объекта по обоим каналам примерно одинаковы, передаточная функция управляющего устройства оказывается физически осуществимой. При наличии не менее двух каналов передачи воздействия f условия абсолютной инвариантности реализуемы (принцип двухканальности Б. Н. Петрова). Подразумевается, что передача одного из каналов выбрана надлежащим образом: в рассмотренном случае — это передаточная функция параллельной компенсации W_{uf} , реализуемой в управляющем устройстве.

Более простые для реализации условия обеспечивают селективную инвариантность. Если, для примера, возмущение имеет постоян-

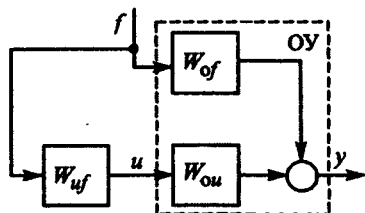


Рис. 3.14. Параллельная компенсация

ные значения, то установившаяся реакция объекта и системы равна нулю при

$$W_{of}(0) + W_{om}(0)W_{uf}(0) = 0$$

или

$$W_{uf}(0) = -W_{of}(0) / W_{om}(0).$$

При этом передаточная функция системы от входа f до выхода y будет иметь нулевой нуль без включения звена с дифференцирующим свойством, как этого требовала последовательная компенсация.

Проблема реализации условий инвариантности в системе с параллельной компенсацией связана с непосредственным измерением возмущения. Это проблема получения полной текущей информации о причинах отклонения управляемой переменной.

3.7.3. Инвариантность в системах с обратной связью

Рассмотрим с позиции обеспечения условий инвариантности управляемой переменной систему с обратной связью (рис. 3.15, а). На рис. 3.15, б изображена типовая структурная схема, отражающая реализацию принципа отрицательной обратной связи в управлении.

Передаточная функция системы от входа f к выходу y запишется так:

$$\Phi_{yf} = \frac{W_{of}}{1 + W_{ou}W_{yy}}. \quad (3.14)$$

Абсолютная инвариантность управляемой переменной достигается при тождественности нулю передаточной функции (3.14), что в условиях

$$W_{of}(s) \neq 0; \quad |W_{of}(s)| < \infty$$

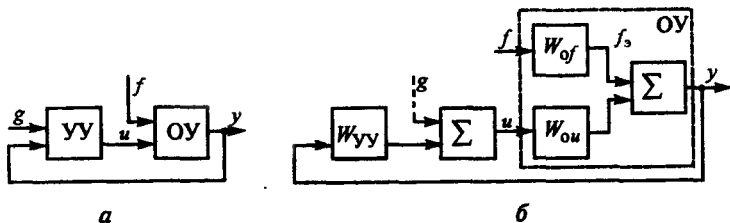


Рис. 3.15. Система с обратной связью

обеспечивается за счет бесконечного усиления контура на всех частотах

$$|W_{ou}(s)W_{yy}(s)| \rightarrow \infty.$$

Мы уже знаем, что при повышении усиления контура возникает проблема устойчивости замкнутой системы. Условие абсолютной инвариантности только за счет обратной связи не может быть реализовано. Отметим, что в рассматриваемой структуре (см. рис. 3.15) имеется единственный канал передачи воздействия.

При конечном усилении и отрицательной обратной связи, как следует из формулы (3.14), происходит ослабление влияния возмущения на выход системы. Можно сказать, не оговаривая величину ϵ , что системы с обратной связью инвариантны до ϵ к возмущениям.

Селективная инвариантность выхода системы с отрицательной обратной связью к возмущениям со спектром $\{s_k\}$ имеет место, если передаточная функция замкнутой системы (3.14) имеет такие же нули, т. е. обращается в нуль на спектре возмущений.

Найдем, какими должны быть передаточные функции объекта по каналу управления $W_{ou} = B_{ou}/A_o$ и управляющего устройства $W_{yy} = B_{yy}/A_{yy}$, чтобы система была селективно инвариантной. Напишем выражение для передаточной функции замкнутой системы (3.14) в виде отношения полиномов. Пусть передаточные функции объекта по обоим каналам имеют одинаковые знаменатели, т. е. $W_{of} = B_{of}/A_o$, тогда

$$\Phi_{yf} = \frac{B_{of}A_{yy}}{A_oA_{yy} + B_{ou}B_{yy}} = \frac{B_{yf}}{A}.$$

Пусть передаточные функции объекта по каналам возмущения и управления не имеют нулей и полюсов, равных полюсам изображения возмущения. Тогда из полученного выражения следует вывод: для селективной инвариантности системы передаточная функция управляющего устройства должна иметь полюсы, равные полюсам изображения возмущения.

Если в практически важном частном случае возмущение имеет постоянные значения, т. е. его изображение имеет нулевой полюс, то необходимо, чтобы $A_{yy}(0) = 0$ — управляющее устройство должно содержать интегратор. Это называют *астатизмом*.

При отсутствии интегратора можно достичь только селективной инвариантности до ϵ — в системе будет иметь место статическая ошибка.

Если воздействие степенное, т. е. его изображение имеет многократный нулевой полюс, то для селективной абсолютной инвариантности

управляющее устройство должно иметь нулевые полюсы той же кратности. В том случае, когда кратность полюса передаточной функции управляющего устройства на единицу меньше кратности полюса возмущения, имеет место селективная инвариантность до ϵ — ненулевая установившаяся ошибка.

Пусть к объекту приложено гармоническое возмущение известной частоты ω_f . Амплитуда установившихся колебаний на выходе замкнутой системы будет равной нулю, если передаточная функция управляющего устройства имеет мнимые полюсы $\pm j\omega_f$, т. е. управляющее устройство содержит консервативное звено.

Таким образом, передаточная функция управляющего устройства должна иметь полюсы, равные полюсам возмущения, т. е. быть «похожей» на изображение возмущения. Полученный результат обобщен в виде *принципа внутренней модели*: в хорошей системе должна присутствовать модель среды [79].

Отметим, что в селективно абсолютно инвариантной системе нулевая установившаяся ошибка сохраняется, независимо от уровня (амплитуды) воздействия, а также от ряда параметров системы. Свойство селективной абсолютной инвариантности, как говорят, *робастно*. Такое свойство достигается благодаря бесконечному усилению контура на комплексных частотах возмущения.

Если усиление контура на частотах возмущения конечно, то имеет место селективная инвариантность до ϵ — конечная установившаяся реакция при ограниченных воздействиях с данным спектром. Ясно, что эта реакция тем меньше, чем больше усиление контура на частотах воздействия.

Пусть гармоническое возмущение, приложенное непосредственно к выходу объекта, имеет частоту ω_f . Для селективной инвариантности до ϵ должно выполняться условие

$$\frac{1}{|1 + W_p(j\omega_f)|} < \epsilon, \quad (3.15)$$

где W_p — передаточная функция разомкнутого контура, равная произведению передаточных функций объекта и управляющего устройства. Поскольку $\epsilon \ll 1$, то

$$|W_p(j\omega)| \gg 1,$$

и вместо (3.15) можно записать

$$\frac{1}{|W_p(j\omega_f)|} < \epsilon$$

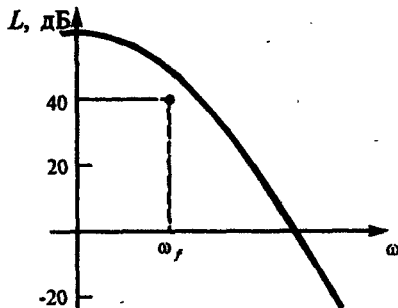


Рис. 3.16. ЛАЧХ разомкнутого контура

или

$$|W_p(j\omega_f)| > 1/\varepsilon.$$

Это значит, что на частоте возмущения усиление контура должно быть достаточно большим. Если использовать логарифмические частотные характеристики, последнее условие примет вид

$$L_p(\omega_f) > -20 \lg \varepsilon.$$

При необходимости ослабить амплитуду возмущения в сто раз ($\varepsilon = 0,01$) усиление контура на частоте возмущения должно быть равно

$$L_p(\omega_f) = -20 \lg 0,01 = 40 \text{ дБ}.$$

На рис. 3.16 изображена ЛАЧХ разомкнутого контура системы, способной ослабить возмущения с частотами $0 < \omega < \omega_f$ не менее чем в сто раз.

3.8. ЧУВСТВИТЕЛЬНОСТЬ СИСТЕМ УПРАВЛЕНИЯ

Воздействия среды на систему до сих пор представлялись как сигналы, преобразуемые операторами звеньев, моделирующих объект и другие элементы системы, или операторами системы в целом. Однако влияние среды на систему может быть также в виде изменения самих операторов преобразования. Отражение в моделях таких изменений приводит к *нестационарным* системам, представляемым, например, в форме дифференциальных уравнений с переменными коэффициентами.

Если изменения операторов во времени много медленнее собственных и вынужденных процессов, то говорят о *квазистационарности*. В этом случае можно рассматривать множество стационарных моделей,

различающихся значениями параметров, а иногда и структурой операторов. Когда на исследуемом отрезке времени известны только номинальные значения и/или интервалы значений параметров, то обычно рассматривают множество стационарных моделей второго ранга $M(2)$, содержащих параметрическую, иначе, структурированную неопределенность.

Для систем управления важно, чтобы малые вариации операторов звеньев не приводили к большим (качественным) изменениям свойств систем в целом, например, к потере устойчивости. Иными словами, необходимо, чтобы система была *грубой*. Работоспособная система управления должна быть инвариантной к возмущениям и устойчивой. Эти свойства системы должны быть также малочувствительны к вариациям операторов звеньев.

Чувствительность количественно характеризует влияние малых изменений свойств элементов на свойства системы. Если изменяются параметры системы, то анализируют чувствительность к этим параметрам. Если малым *неструктурированным* вариациям подвержены операторы звеньев систем с раскрытой причинно-следственной структурой, то применяют методы структурной теории чувствительности. Чувствительность передаточных функций систем позволяет анализировать влияние свойств звеньев на условия инвариантности и ковариантности, а чувствительность характеристических полиномов — на условия устойчивости и характер переходных процессов [16, 67, 76, 88].

Абсолютная чувствительность передаточной функции системы $\Phi(s)$ к вариации передаточной функции звена $W(s)$ определяется как частная производная

$$T_W^\Phi = \partial \Phi(W) / \partial W.$$

Эта функция комплексного аргумента s позволяет в первом приближении найти вариацию передаточной функции системы по известной вариации передаточной функции звена

$$\delta \Phi(s) \approx T_W^\Phi(s) \delta W(s).$$

Относительная чувствительность S_W^Φ отражает связь между относительными вариациями передаточных функций или частотных характеристик

$$\frac{\delta \Phi}{\Phi} \approx S_W^\Phi \frac{\delta W}{W}.$$

Связь между относительной и абсолютной функциями чувствительности имеет вид

$$S_W^\Phi = T_W^\Phi \frac{W}{\Phi}. \quad (3.16)$$

Функции чувствительности — рациональные функции комплексного аргумента. Говоря о величине функции чувствительности, имеют в виду ее модуль. Если вместо s подставить $j\omega$, то по функции чувствительности можно найти вариации амплитудных частотных характеристик. В частном случае при $s = 0$ в приведенных соотношениях фигурируют действительные числа — коэффициенты усиления.

3.9. ЧУВСТВИТЕЛЬНОСТЬ СИСТЕМ С ТИПОВОЙ СТРУКТУРОЙ

Чувствительность характеристик системы, образованной соединением звеньев, зависит от ее структуры. Рассмотрим вначале системы с типовыми структурами.

3.9.1. Чувствительность системы, образованной последовательным соединением звеньев

Передаточная функция системы, образованной последовательным соединением звеньев (см. рис. 2.9, а), равна произведению передаточных функций

$$\Phi = W_1 W_2.$$

Пусть изменяются характеристики первого звена

$$W_1 = W_{10} + \delta W_1.$$

Передаточная функция соединения при этом запишется так:

$$\Phi = \Phi_0 + \delta\Phi = (W_{10} + \delta W_1) W_2 = W_{10} W_2 + W_2 \delta W_1,$$

т. е. абсолютная функция чувствительности

$$T_{W_1}^\Phi = W_2.$$

Относительная чувствительность определяется по формуле (3.16):

$$S_{W_1}^\Phi = 1.$$

Тот же результат получится и для относительной чувствительности к W_2 .

Вообще, относительная чувствительность последовательного соединения к любому из звеньев равна единице. Это значит, что изменение модуля передаточной функции или АЧХ любого из звеньев на сколько-то процентов приводит к изменению модуля передаточной функции системы и ее частотной характеристики на столько же процентов. Следовательно, нельзя изменить относительную чувствительность системы путем последовательного включения новых звеньев или удаления части существующих.

3.9.2. Чувствительность системы, образованной параллельным соединением звеньев

Передаточная функция системы, образованной параллельным соединением звеньев (рис. 2.9, б), равна сумме передаточных функций:

$$\Phi = W_1 + W_2.$$

Пусть изменяется первое звено

$$W_1 = W_{10} + \delta W_1.$$

Передаточная функция системы при этом запишется так:

$$\Phi = \Phi_0 + \delta\Phi = W_{10} + \delta W_1 + W_2.$$

Следовательно, абсолютная функция чувствительности для этой структуры всегда равна единице

$$T_{W_1}^\Phi = 1.$$

Выражение для относительной функции чувствительности находится по формуле (3.16):

$$S_{W_1}^\Phi = \frac{W_1}{W_1 + W_2}.$$

Чем больше модуль передаточной функции варьируемого звена, тем больше относительная чувствительность, т. е. вклад звена в передачу системы в целом. Если

$$|W_1(s)| \gg |W_2(s)|,$$

то $S_{W_1}^\Phi \approx 1$; если

$$|W_1(s)| \ll |W_2(s)|,$$

то $S_{W_1}^\Phi \approx 0$.

Уменьшить чувствительность передачи системы к вариациям звена можно увеличением усиления неварьируемого звена, подсоединенного параллельно.

3.9.3. Чувствительность системы с обратной связью

Передаточная функция системы с отрицательной обратной связью (см. рис. 2.9, в) равна

$$\Phi = \frac{W_1}{1 + W_1 W_2}.$$

Пусть варьирует передача звена, находящегося в прямом пути:

$$W_1 = W_{10} + \delta W_1.$$

Абсолютная функция чувствительности находится дифференцированием выражения $\Phi(W_1)$:

$$T_{W_1}^{\Phi} = \frac{1}{(1 + W_1 W_2)^2}.$$

Относительная функция чувствительности определяется по формуле (3.16):

$$S_{W_1}^{\Phi} = \frac{1}{1 + W_1 W_2}, \quad (3.17)$$

она равна отношению передаточной функции системы к передаточной функции варьируемого звена

$$S_{W_1}^{\Phi} = \Phi / W_1.$$

По выражению (3.17) можно заключить, что использование обратной связи позволяет существенно изменить относительную чувствительность. При повышении усиления контура за счет любого из звеньев относительная чувствительность системы с отрицательной обратной связью уменьшается. Отрицательная обратная связь существенно уменьшает влияние вариаций в прямом пути на передачу всего соединения, если велико усиление контура. Этим широко пользуются на практике. Поскольку объект управления находится в прямом пути, то повышение усиления контура при стабильной обратной связи уменьшает чувствительность передачи системы по каналу воспроизведения задающего воздействия.

В динамических системах усиление контура на различных частотах неодинаково. Следовательно, обратная связь уменьшает чувствительность только в тех интервалах частот, где усиление велико. На частотах, где усиление контура мало, относительная чувствительность характеристики замкнутой системы будет близка к единице.

В п. 3.7.3 выяснилось, что большое усиление контура обеспечивает инвариантность переменной выхода системы с обратной связью к сигнальному возмущению. Далее установлено, что одновременно ослабляется и влияние параметрических воздействий среды. В этом заключается универсальность обратной связи.

Пусть теперь изменяются характеристики звена обратной связи:

$$W_2 = W_{20} + \delta W_2.$$

Абсолютная функция чувствительности передачи системы к вариациям звена

$$T_{W_2}^{\Phi} = -\frac{W_1^2}{(1 + W_1 W_2)^2}.$$

Выражение для относительной чувствительности имеет вид

$$S_{W_2}^{\Phi} = -\frac{W_1 W_2}{1 + W_1 W_2}; \quad (3.18)$$

она равна произведению передаточной функции системы и передаточной функции варьируемого звена с обратным знаком:

$$S_{W_2}^{\Phi} = -\Phi W_2.$$

При повышении усиления контура относительная чувствительность увеличивается

$$|W_1(s)W_2(s)| \rightarrow \infty, \quad S_{W_2}^{\Phi} \rightarrow -1.$$

Применительно к системам управления, реализующим принцип управления по отклонению, можно говорить о необходимости стабильной обратной связи.

3.10. ЧУВСТВИТЕЛЬНОСТЬ СИСТЕМ СО СЛОЖНОЙ СТРУКТУРОЙ

В случае систем произвольной структуры для анализа чувствительности целесообразно использовать частотный метод структурной теории чувствительности [88]. Приведем некоторые положения этой теории.

Относительные функции чувствительности системы к неструктурированным вариациям звеньев могут быть вычислены как передаточные функции некоторой преобразованной системы. Преобразования сводятся к введению в графы единичных дуг, так называемых *диполей чувствительности*. Положение диполя чувствительности в графе зависит от *позиции* варьируемой дуги. Пусть имеется единственный путь в графе. Дуга находится в позиции 1, если принадлежит пути. Дуга находится в позиции 2, если принадлежит контурам, касающимся пути, но в путь не входит. Наконец, позиция 3 соответствует ситуации, когда дуга входит в контуры, не касающиеся пути. Вложение диполя чувствительности в дугу означает последовательное соединение, иначе, под-разбиение варьируемой дуги.

Если варьируемая дуга находится в позиции 1, то функция чувствительности равна передаточной функции графа от вершины-входа дипо-

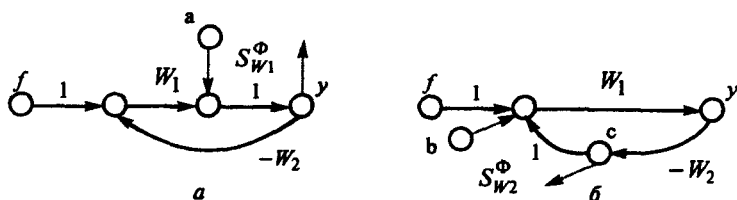


Рис. 3.17. Графы с диполями чувствительности:
а — прямым; б — обратным

для чувствительности, вложенного в варьируемую дугу, до его вершины-выхода. Найдем выражение для функции чувствительности системы с обратной связью (см. рис. 2.9, в) к дуге W_1 , находящейся в позиции 1 к пути. Вложим диполь чувствительности (рис. 3.17, а) и вычислим передаточную функцию $\Phi_{yu}(s)$ от входа до выхода диполя. Результат совпадает с выражением (3.17), полученным в п. 3.9.3 дифференцированием передаточной функции.

Если варьируемая дуга находится в позиции 2, то функция чувствительности равна передаточной функции от вершины-выхода до вершины-входа единичной дуги (обратный диполь чувствительности). Для рассматриваемой системы соответствующий граф с обратным диполем чувствительности показан на рис. 3.17, б. Выражение для функции чувствительности $S_{W_2}^{\Phi} = \Phi_{cb}$ совпадает с полученным в п. 3.9.3 выражением (3.18).

3.11. ПОКАЗАТЕЛИ КАЧЕСТВА СОБСТВЕННЫХ ДВИЖЕНИЙ СИСТЕМ УПРАВЛЕНИЯ

Показатели качества собственно систем управления M_s (иначе, автономных систем) количественно оценивают свободные движения, которые являются решениями однородных дифференциальных уравнений при произвольных предначальных условиях:

$$A(p)x(t) = 0; \quad x(0_-), \quad x'(0_-), \dots, x^{(n-1)}(0_-).$$

Необходимым условием работоспособности системы управления является затухание собственных движений при любых начальных условиях, т. е. устойчивость системы. Однако это условие не является достаточным — затухание переходных процессов должно быть достаточно быстрым, они не должны быть чрезмерно колебательными. Говорят, что система должна иметь запас устойчивости (или грубости).

Изображение свободных движений по Лапласу имеет вид

$$X(s) = A_n(s)/A(s),$$

где

$$A(s) = a_0 + a_1 s + \dots + a_{n-1} s^{n-1} + a_n s^n.$$

Полином $A_n(s)$, учитывающий ненулевые предначальные условия, может быть записан в виде

$$A_n(s) = M(s)x(0_-),$$

где $x(0_-) = (x(0_-), x'(0_-), \dots, x^{(n-1)}(0_-))^T$ — вектор начальных условий; $M(s) = (A_{n-1} \dots A_1, A_0)$ — полиномиальная строчная матрица с элементами

$$A_0 = a_n;$$

$$A_1 = a_n s + a_{n-1};$$

.....

$$A_{n-1} = a_n s^{n-1} + \dots + a_2 s + a_1.$$

Как показано ранее, выражение для свободных движений в случае простых корней характеристического полинома запишется так:

$$x(t) = \sum_{i=1}^n \frac{A_n(s_i)}{A'(s_i)} e^{s_i t}.$$

Свободное движение определяется корнями характеристического полинома и начальными условиями.

Для автономных систем предначальные условия неопределенные. Поэтому здесь ведущую роль играют показатели качества, характеризующие области расположения корней. Вместе с тем, полезное образное представление о характере свободных движений дают кривые решения однородных дифференциальных уравнений при типовых начальных условиях:

$$x(0_-) = (1 \ 0 \dots 0)^T,$$

когда выражение для свободного движения принимает вид

$$x(t) = \sum_{i=1}^n \frac{a_n s_i^{n-1} + \dots + a_2 s_i + a_1}{n a_n s_i^{n-1} + \dots + 2 a_2 s_i + a_1} e^{s_i t}.$$

На рис. 3.18, а изображены кривые монотонного (1), аperiodического (2) и колебательного (3) процессов, а на рис. 3.18, б представлены прямые показатели качества свободного движения: t_p — время регулирования; $x_{\max 1}$ — максимальное отклонение.

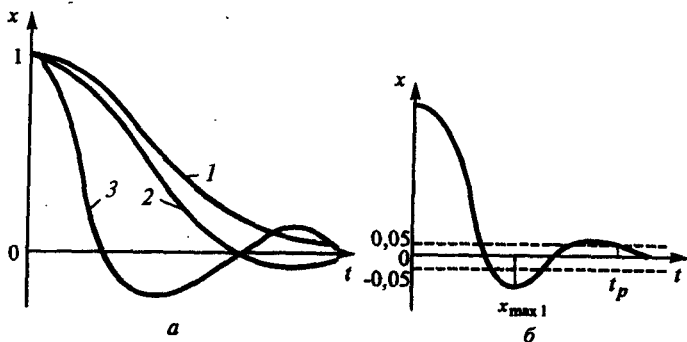


Рис. 3.18. Кривые, отражающие различный характер процессов (а) и определение прямых показателей качества (б)

3.11.1. Корневые показатели качества процессов управления

Выражение для собственных движений при типовых начальных условиях позволяет сформулировать некоторые требования к расположению корней характеристического полинома на комплексной плоскости, при выполнении которых имеют место определенный тип и параметры процессов. Эти условия называются *корневыми показателями качества*. Таким образом, определенные суждения о процессах можно высказать и в том случае, когда значения корней не вычисляются, но имеется некоторая информация об их расположении. Например, установление факта принадлежности всех корней левой полуплоскости равносильно знанию об асимптотическом затухании процессов при любых начальных условиях. Поскольку многие показатели качества имеют смысл лишь для устойчивых процессов, именно с определения устойчивости начинают анализ качества.

Связь между прямыми показателями качества и расположением корней на комплексной плоскости может выражаться как соответствие между параметрами некоторой области Γ на комплексной плоскости, где расположены все корни, и области Λ на плоскости временных характеристик (t, x) , где расположены решения однородного дифференциального уравнения при типовых начальных условиях. Такая связь в общем виде, во-первых, достаточно сложна, во-вторых, область Λ получается слишком обширной, что ведет к грубой оценке прямых показателей качества. Поэтому соответствие областей Γ и Λ устанавливается в зависимости от классов распределения корней [82]: класс 0 — все

корни действительные; класс 1 — среди корней имеется пара комплексных; класс 2 — среди корней имеются две пары комплексных и т. д. Кроме того, различают подклассы: А — ближайшей к мнимой оси является пара комплексных корней; Б — ближайшим является вещественный корень.

Для класса 0 область Λ определяется так:

$$u(t) \leq x(t) \leq v(t),$$

где миноранта

$$u(t) = \exp\{-\tau\},$$

а мажоранта

$$v(\tau) = \left[1 + \tau + \tau^2/2 + \dots + \tau^{n-1}/(n-1)!\right] \exp\{-\tau\}.$$

В этих выражениях введено относительное время $\tau = \eta t$ (η — расстояние от мнимой оси до ближайшего к ней корня). На рис. 3.19, а приведены кривые $v(\tau)$ для $n = 1, \dots, 6$. Очевидно, для $n = 1$ имеем: $v(\tau) = u(\tau)$.

Расстояние η до ближайшего к мнимой оси корня называют *степенью устойчивости*. Ясно, что чем больше η , тем быстрее затухают апериодические переходные процессы (больше абсолютное затухание). Область Γ для класса 0 представляет собой луч на вещественной оси (рис. 3.19, б).

Если среди корней имеется хотя бы одна пара комплексных (класс 1), то область Λ определяется так:

$$u_1(t) \leq x(t) \leq v_1(t),$$

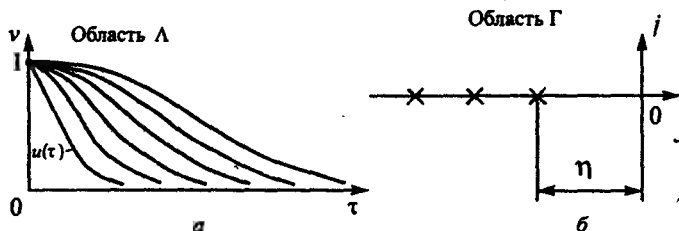


Рис. 3.19. Области для случая действительных корней: а — временная плоскость; б — комплексная плоскость

где мажоранта та же, что и в классе 0, а минорантой является кусочная кривая, описываемая уравнением;

$$u_1(\tau) = \begin{cases} \exp\{-\tau\}(a \exp\{-b\tau\} + 1 - a), & 0 \leq \tau \leq \tau_0; \\ -\exp\{-\pi/\mu\}, & \tau_0 < \tau \leq \tau_1; \\ -v_1(\tau), & \tau > \tau_1, \end{cases}$$

где $a = \exp\{b\tau_0\}/b$; $2b = \exp\{b\tau_0\} - 1$. Коэффициенты a, b выбираются таким образом, чтобы кривая, соответствующая уравнению для $u_1(\tau)$ на интервале $0 \leq \tau \leq \tau_0$, имела горизонтальную касательную в точке $\tau = \tau_0$, где $u_1 = -\exp\{-\pi/\mu\}$. Точка τ_1 определяется как показано на рис. 3.20, а. Соответствующая область Γ , характеризуемая двумя корневыми показателями качества — степенью устойчивости η и колебательностью μ , — изображена на рис. 3.20, б. Чем меньше колебательность μ — отношение мнимой части корня к действительной, тем меньше прямой показатель — максимальное отклонение $x_{\max}$.

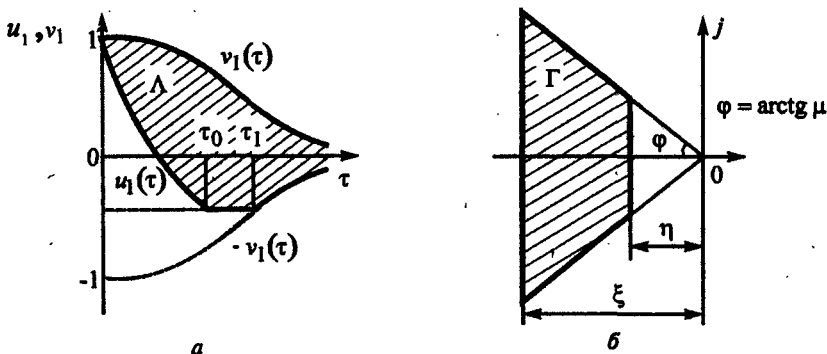


Рис. 3.20. Области на плоскостях для случая пары комплексных корней:
а — временной; б — комплексной

Если кроме показателей качества η и μ ввести еще один, например $\xi = \max |\operatorname{Re} s_i|$ (см. рис. 3.20, б), то область Λ может быть сделана еще более узкой. Следовательно, суждение о процессе станет более определенным. Однако при этом выражения для границ области Λ становятся более сложными и здесь не приводятся.

Иногда используется наиболее простая для вычисления корневая оценка — среднегеометрическое значение корней

$$\Omega = \sqrt[n]{(-1)^n s_1 \dots s_n} = \sqrt[n]{a_0 / a_n},$$

которая может служить относительной мерой быстродействия систем с однотипным распределением корней.

Если требуется лишь определить, обладает ли система некоторой степенью устойчивости η , можно обойтись без непосредственного вычисления корней характеристического полинома. Задача решается построением смещенного полинома

$$A_{\eta}(z) = A(s) \Big|_{s=z-\eta}.$$

Геометрически такая замена аргумента означает смещение оси ординат плоскости корней влево на расстояние η . Применяя к смещенному полиному $A_{\eta}(z)$ критерий Рауса или Гурвица, можно сделать вывод о расположении его корней относительно мнимой оси. Если все корни полинома $A_{\eta}(z)$ левые, то система с характеристическим полиномом $A(s)$ обладает степенью устойчивости η , и наоборот.

Этот прием можно обобщить. Пусть на плоскости s задана область Γ и необходимо определить, принадлежат ли этой области корни полинома $A(s)$. Положим, что нет алгоритма, решающего задачу для области Γ , но имеется алгоритм для другой конфигурации области, например Δ . Вводим отображение $s = \varphi(\delta)$, переводящее Γ в Δ , т.е. такое, что из факта принадлежности корней $A(s)$ области Γ следует принадлежность корней $A_{\Delta}(\delta)$ области Δ , и наоборот. Вычисляем коэффициенты $A_{\Delta}(\delta)$ и применяем соответствующий алгоритм проверки принадлежности корней области Δ . Разумеется, при этом следует иметь алгоритм преобразования коэффициентов полиномов.

3.11.2. Интегральные показатели качества

Интегральные показатели качества характеризуют устойчивые процессы. При этом функции времени $x(t)$ ограничены и с течением времени стремятся к нулю, т.е. абсолютно интегрируемы. Интегральные показатели качества — функционалы, в которых роль независимой переменной играют функции времени $x(t)$; $t \geq 0$:

$$J = \int_0^{\infty} F(x(t)) dt,$$

где функция F подбирается так, чтобы несобственный интеграл сошелся, а значения J количественно характеризовали качество процессов $x(t)$. Часто функционал J представляет собой норму функционального пространства на интервале $[0, \infty)$.

Интегральные показатели качества оценивают процессы в наиболее сжатом виде — кривой $x(t)$ ставится в соответствие число. Обычно, чем меньше значение J , тем процесс сходится быстрее, имеет меньшие значения ординат для определенных t .

Простейший линейный интегральный показатель качества

$$J_0 = \int_0^{\infty} x(t) dt$$

представляет собой площадь под кривой (рис. 3.21). Если процесс не меняет знака, то площадь под кривой тем меньше, чем быстрее затухает процесс.

Обобщением линейного интегрального показателя являются моменты

$$J_l = \int_0^{\infty} t^l x(t) dt, \quad l = 0, 1, \dots,$$

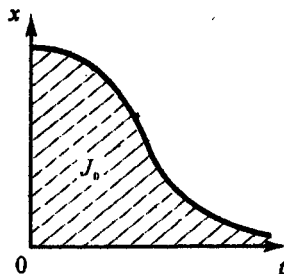


Рис. 3.21. Иллюстрация интегрального показателя качества

для вычисления которых производится интегрирование с весом — больше «стоят» ординаты процесса при больших t . Моменты легко вычисляются по изображению переменной $X(s)$, если воспользоваться теоремой о дифференцировании изображения

$$\int_0^{\infty} t^l x(t) e^{-st} dt = (-1)^l \frac{d^l}{ds^l} X(s).$$

Положим $s = 0$, тогда получим

$$J_l = \int_0^{\infty} t^l x(t) dt = (-1)^l \frac{d^l}{ds^l} X(s) \Big|_{s=0}.$$

В частном случае линейного интегрального показателя имеем:

$$J_0 = X(s) \Big|_{s=0}.$$

Недостаток использования моментов в том, что они характеризуют только процессы, не меняющие знак.

Хорошо оценивает процессы, меняющие знак, интегральный квадратичный показатель

$$I_0 = \int_0^{\infty} x^2(t) dt.$$

Этот показатель тем меньше, чем быстрее процесс приближается к идеальному, для которого $I_0 = 0$.

Преобразуем выражение для интегрального квадратичного показателя с учетом формулы обращения Лапласа

$$I_0 = \int_0^{\infty} x(t)x(t)dt = \int_0^{\infty} x(t) \left(\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} X(s)e^{st}ds \right) dt$$

(в устойчивых системах абсцисса сходимости равна нулю) и поменяем порядок интегрирования

$$I_0 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} X(s) \left[\int_0^{\infty} x(t)e^{st}dt \right] ds.$$

Выражение в квадратных скобках есть $X(-s)$, следовательно, получим:

$$I_0 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} X(s)X(-s)ds.$$

Под интегралом имеем произведение рациональных функций. Для его вычисления можно воспользоваться готовыми формулами (для конкретных степеней n полиномов знаменателей). Например,

для $n=1$

$$X(s) = \frac{b_0}{a_1s + a_0}, \quad I_0 = \frac{b_0^2}{2a_0a_1};$$

для $n=2$

$$X(s) = \frac{b_1s + b_0}{a_2s^2 + a_1s + a_0}; \quad I_0 = \frac{b_1^2a_0 + b_0^2a_2}{2a_0a_1a_2};$$

для $n=3$

$$X(s) = \frac{b_2^2s^2 + b_1s + b_0}{a_3s^3 + a_2s^2 + a_1s + a_0},$$

$$I_0 = \frac{b_2^2a_0a_1 + (b_1^2 - 2b_0b_2)a_0a_3 + b_0^2a_2a_3}{2a_0a_3(a_1a_2 - a_0a_3)}.$$

В общем случае для вычисления интегральных квадратичных показателей целесообразно использовать алгоритм Некольна — Острема (I. Nkolny — K. J. Åström) [55].

Если положить $s = j\omega$, то получим:

$$I_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega)X(-j\omega)d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(j\omega)|^2 d\omega.$$

Это соотношение называют формулой Парсеваля (M. Parseval). Оно означает, что энергия движения пропорциональна интегралу от квадрата модуля спектральной функции этого процесса.

Недостатком интегрального квадратичного показателя является то, что он не учитывает колебательности процесса. Во многих случаях системы с меньшим значением этого показателя оказываются более колебательными.

Пример 2. Рассмотрим систему с единичной отрицательной обратной связью, передаточная функция разомкнутого контура которой равна

$$W_p(s) = \frac{k}{s(s+1)}.$$

Передаточная функция замкнутой системы от задающего воздействия до переменной ошибки e записывается так:

$$W_e(s) = \frac{1}{1 + W_p(s)} = \frac{s(s+1)}{s(s+1) + k}.$$

Изображение переменной ошибки $E(s)$ в случае задания в виде единичного ступенчатого воздействия равно

$$E(s) = W_e(s)/s = \frac{s+1}{s^2 + s + k}.$$

Найдем выражение интегральной квадратичной ошибки по приведенной ранее формуле для $n = 2$:

$$I_0 = (k+1)/2k = 0,5(1+1/k).$$

Как видно из полученной зависимости $I_0(k)$, значение показателя качества уменьшается с увеличением коэффициента усиления контура; при $k \rightarrow \infty$ имеем предельное значение: $I_0 \rightarrow 1/2$.

Рассматриваемая система второго порядка устойчива при любом положительном значении k . Действительно, при этом положительны все коэффициенты характеристического полинома замкнутой системы

$$s^2 + s + k,$$

для полинома второй степени это необходимое условие устойчивости является и достаточным (см. п. 3.3.1).

Однако стремление к уменьшению интегрального квадратичного показателя качества путем увеличения коэффициента усиления контура вызывает рост колебательности процессов. Чтобы убедиться в этом, получим выражение для колебательности корней μ , который равен отношению абсолютных значений мнимой и действительной частей корней (см. п. 3.11.1). Из выражения для корней характеристического полинома

$$s_{1,2} = (-1 \pm \sqrt{4k-1})/2$$

видно, что при $k > 1/4$ корни становятся комплексными, а процессы — колебательными. Для этих значений коэффициента усиления имеем следующее выражение для корневого показателя качества — колебательности корней:

$$\mu = \sqrt{4k-1}.$$

С увеличением k растет колебательность корней μ , а значит, уменьшается относительное затухание (затухание за период) процессов. Заметим, что при этом абсолютное затухание, определяемое действительной частью корней, остается неизменным.

Таким образом, попытка минимизации интегрального квадратичного показателя качества даже для системы второго порядка приводит к чрезмерно колебательным процессам. Это можно объяснить тем, что экстремалью функционала I_0 является идеальный процесс, затухающий мгновенно; стремление к такому нереальному в динамических системах процессу и дает приведенный в примере результат.

Для учета колебательности процесса интегральный квадратичный показатель дополняют взвешенным интегралом от квадрата производной $x'(t)$, т. е. переходят к выражению

$$I_1 = \int_0^{\infty} [x^2(t) + \tau_1^2 (x')^2(t)] dt.$$

Этот показатель называется *улучшенным интегральным квадратичным показателем*. Если процесс колебательный, то относительно большие значения имеют ординаты $x'(t)$, что и учитывается интегралом с весовым коэффициентом τ_1^2 .

Улучшенная интегральная квадратичная оценка тем меньше, чем ближе процесс $x(t)$ к некоторой экстремали $x^*(t)$. Найдем эту кривую, играющую роль, аналогичную экстремуму функции, т. е. найдем кривую, доставляющую минимум функционалу. Для этого преобразуем I_1 к виду

$$I_1 = \int_0^{\infty} [x(t) + \tau_1 x'(t)]^2 dt - 2\tau_1 \int_0^{\infty} x(t)x'(t) dt.$$

Второе слагаемое запишем так:

$$-2\tau_1 \int_{x(0)}^0 x dx = -2\tau_1 \frac{x^2}{2} \Big|_{x(0)}^0 = \tau_1 x^2(0) = \tau_1.$$

Здесь учтено, что $x(\infty) = 0$, а $x(0) = 1$.

Функционал I_1 достигает минимума, если первая составляющая I_1 равна нулю, что означает

$$\tau_1 x'(t) + x(t) = 0.$$

Следовательно, I_1 достигает минимума, если $x(t)$ является решением дифференциального уравнения первого порядка при начальном условии $x(0)$, т. е. когда

$$x(t) = x^*(t) = x(0)e^{-t/\tau_1}.$$

Эта кривая — экстремаль функционала I_1 — изображена на рис. 3.22. Минимальное значение функционала при типовом начальном условии $x(0) = 1$ равно

$$I_1(x^*(t)) = \tau_1.$$

Чем лучше система управления в смысле показателя I_1 , тем ближе процесс к экспоненте $x^*(t)$ с постоянной времени τ_1 . Увеличение коэффициента веса τ_1 при формировании показателя означает предпочтение системам с более медленно затухающими процессами. По существу, показатель I_1 означает принятие стереотипа качества в виде эталонной системы первого порядка. Это более реальное (по сравнению с интегральным квадратичным показателем) пожелание к поведению динамических систем.

Разумеется, в случае, если анализируется система высокого порядка процесс $x(t)$ отличается от экспоненты $x^*(t)$. При этом разность

$$\delta I_1 = I_1(x(t)) - I_1(x^*(t)) = I_1 - \tau_1 \geq 0$$

имеет смысл расстояния между кривыми $x(t)$, $x^*(t)$.

Однако по величине δI_1 не так удобно судить, насколько прямые показатели процесса $x(t)$ отличаются от соответствующих показателей экспоненты. Гораздо удобнее, если иметь оценку

$$\max_t |x(t) - x^*(t)|$$

— максимум модуля разности кривых. Целесообразно найти по δI_1 мажоранту $x^*(t) + u$ и миноранту $x^*(t) - u$ для $x(t)$. Оказывается [82], что для улучшенной интегральной квадратичной оценки справедливо соотношение:

$$u = \sqrt{\delta I_1 / \tau_1}.$$

На рис. 3.23 изображены мажоранта и миноранта. Чем больше величина τ_1 , тем уже полоса, в которую заключена оцениваемая кривая $x(t)$, однако больше и значение функционала I_1 . Учитывая, что минимальное значение I_1 равно τ_1 , запишем

$$u = \sqrt{(I_1 - \tau_1) / \tau_1} = \sqrt{(I_1 / \tau_1) - 1}.$$

Наилучшая оценка кривой $x(t)$ по улучшенному интегральному показателю качества достигается при минимальном значении u , которое зависит от значения коэффициента веса τ_1 . Принимая различные τ_1 , можно добиться наилучшей оценки прямых показателей качества процесса $x(t)$ через интегральный показатель качества $I_1(x(t), \tau_1)$.

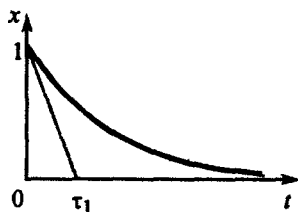


Рис. 3.22. Безусловная экстремаль улучшенного интегрального показателя

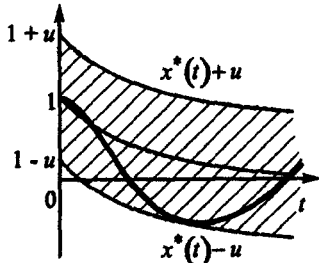


Рис. 3.23. Мажоранта и миноранта процессов

При необходимости оценки процессов путем сопоставления с эталонным поведением, описываемым дифференциальным уравнением n -го порядка, привлекают обобщенные интегральные квадратичные показатели вида

$$I_n = \int_0^{\infty} \left[x^2 + \tau_1^2 (\dot{x})^2 + \dots + \tau_n^{2n} (x^{(n)})^2 \right] dt. \quad (3.19)$$

Если модель собственно системы управления представлена в нормальной форме пространства состояний

$$\frac{dv}{dt} = Av; \quad v(0),$$

при типовых начальных условиях $v(0) = (1 \ 0 \ \dots \ 0)^T$, то обобщенная интегральная оценка запишется так:

$$I_n = \int_0^{\infty} v^T Q v dt, \quad (3.20)$$

где Q — положительно определенная (неотрицательно определенная) матрица весовых коэффициентов. Подынтегральное выражение представляет собой квадратичную форму. Например, для $n=2$ имеем

$$v^T Q v = (v_1 v_2) \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = q_{11} v_1^2 + q_{12} v_1 v_2 + q_{21} v_2 v_1 + q_{22} v_2^2.$$

Выражение (3.19) для обобщенного интегрального показателя является частным случаем (3.20), когда $v_1 = x$, а матрица Q — диагональная (квадратичная форма — каноническая).

3.11.3. Частотные показатели качества

Частотные показатели качества свободных движений систем с обратной связью базируются на частотных характеристиках разомкнутых контуров $W_p(j\omega)$. Пусть одноконтурная система образована устойчивыми звеньями. Замыкание контура и повышение его усиления могут привести к тому, что некоторые корни характеристического полинома $A(s)$ системы будут приближаться к границе устойчивости — мнимой оси. Если АФХ $W_p(j\omega)$ проходит на частоте ω_0 через точку $(-1, j0)$, то характеристический полином замкнутой системы $A(s)$ имеет пару мнимых корней $\pm j\omega_0$. Если же годограф $W_p(j\omega)$ не охватывает критическую точку и проходит на некотором удалении от нее, то замкнутая система будет иметь запас устойчивости.

Для количественной оценки удаленности годографа $W_p(j\omega)$ от точки $(-1, j0)$, а следовательно, левых корней $A(s)$ от мнимой оси, вводят понятия *запаса устойчивости по модулю* (амплитуде) и *запаса устойчивости по фазе*.

На рис. 3.24, а изображена типовая АФХ устойчивой разомкнутой системы, которая будет устойчивой и после замыкания контура. Запас устойчивости по модулю β определяется как отношение

$$\beta = 1/|W_p(j\omega_\pi)|,$$

где ω_π — частота, на которой фазовый сдвиг, вносимый контуром, равен $-\pi$ рад. Для устойчивых систем запас по модулю больше единицы. Если система находится на границе устойчивости, то

$$W_p(j\omega_\pi) = -1; \quad \beta = 1.$$

Величина β обычно лежит в пределах 2...10. Запас устойчивости по фазе $\Delta\phi$ определяется так:

$$\Delta\phi = \pi + \arg W_p(j\omega_{\text{ср}}),$$

где $\omega_{\text{ср}}$ — частота среза, на которой модуль АФХ разомкнутого контура равен единице. Величина запаса устойчивости по фазе обычно лежит в пределах $30^\circ \dots 60^\circ$.

Запас по модулю удобно выражать в логарифмическом масштабе

$$\Delta L = 20 \lg \beta = -20 \lg |W_p(j\omega_\pi)| = -L_p(\omega_\pi).$$

Его величина обычно лежит в пределах 6 ... 20 дБ. На рис. 3.24, б показаны запасы устойчивости по модулю и по фазе, определяемые по логарифмическим частотным характеристикам.

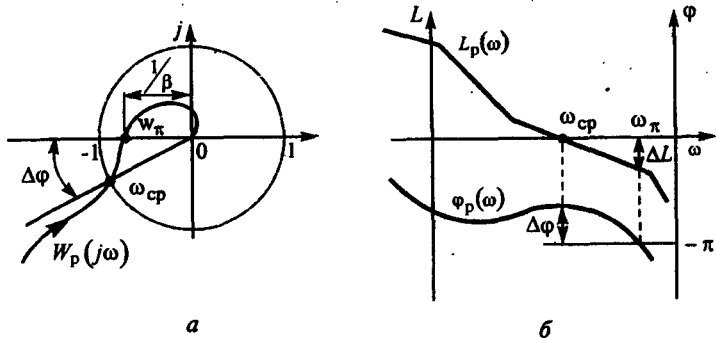


Рис. 3.24. Частотные показатели качества

Показатели качества — запас по модулю ΔL и запас по фазе $\Delta\phi$ — количественно характеризуют роль контура в перемещении корней характеристического полинома, как правило, доминирующих, т. е. ближайших к мнимой оси комплексных корней, в основном определяющих характер процессов.

Частота среза ω_{cp} также является косвенным показателем качества процессов, характеризую их быстродействие. Эта частота приблизительно равна мнимой части доминирующих комплексных корней, а следовательно, частоте колебаний процессов. Если положить, что за время t_p имеет место 1...2 колебания, то

$$t_p \approx \frac{(1...2)2\pi}{\omega_{cp}}.$$

Связи между параметрами типовых логарифмических характеристик разомкнутых систем $L_p(\omega)$ и прямыми, а также некоторыми косвенными показателями качества замкнутых систем хорошо изучены. Имеются специальные номограммы [64, 90], с помощью которых для различных типов асимптотических логарифмических характеристик можно определить показатели качества замкнутых систем. Вместе с тем в условиях применения ЭВМ для расчетов систем управления роль этих номограмм снижается.

3.12. ПОКАЗАТЕЛИ КАЧЕСТВА ВЫНУЖДЕННЫХ ПРОЦЕССОВ УПРАВЛЕНИЯ

Рассмотрим вопросы анализа систем управления по моделям типа M_{YSF} , получающимся дополнением моделей собственно систем управ-

ления M_s информацией о точках приложения воздействий на входе и точках выхода, т. е. указанием переменных систем, принятых за входы и выходы (см. рис. 1.18). О воздействиях при этом нет никакой информации, т. е. модели этого типа характеризуют только каналы преобразования воздействий. Модели M_{TSF} представляются в форме совокупности неоднородных дифференциальных уравнений (2.1), в форме передаточных функций (2.3), дифференциальными уравнениями в форме пространства состояний

$$\frac{dv}{dt} = Av + Bf;$$

$$y = Cv + Df,$$

а также частотными или временными характеристиками.

В системах управления выделяют два типа каналов передач воздействий:

1 — каналы подавления возмущений $f_d(t)$; $d = 1, 2, \dots$, с передаточными функциями

$$\Phi_{fd}(s) = B_{fd}(s)/A(s);$$

2 — канал воспроизведения задания $g(t)$ с передаточной функцией

$$\Phi_g(s) = B_g(s)/A(s).$$

Передаточная функция идеального канала подавления возмущения тождественна нулю

$$\Phi_{fd}(s) \equiv 0,$$

что соответствует абсолютной инвариантности, а передаточная функция идеального канала воспроизведения задания тождественна единице

$$\Phi_g(s) \equiv 1,$$

что означает абсолютную ковариантность. Реально к передачам предъявляют требования селективной абсолютной инвариантности и ковариантности, а также инвариантности и ковариантности до ϵ .

Прямые показатели качества процессов используются при типовых воздействиях и нулевых предначальных условиях, в частности, количественно оценивают переходные характеристики $h(t)$. На рис. 3.25 изображена типовая переходная характеристика системы управления по каналу воспроизведения. Прямыми показателями качества являются: $y_{уст}$ — установившееся значение процесса; t_p — время регулирования, определяемое как время, по истечении которого процесс не выходит за пределы 5% зоны в окрестности установившегося значения; перерегулирование

$$\sigma = \frac{y_{\max} - y_{\text{уст}}}{y_{\text{уст}}} 100\%.$$

При рассмотрении канала воспроизведения установившееся значение выхода системы — управляемой переменной — равно единице или близко к ней. По существу, отличное от единицы значение $y_{\text{уст}}$ количественно характеризует селективную ковариантность до ϵ выхода с постоянным заданием.

На рис. 3.26 изображена типовая переходная характеристика по каналу подавления возмущения; там же указаны прямые показатели качества. Здесь установившееся значение выхода системы близко к нулю. Это количественная характеристика селективной инвариантности до ϵ выхода к постоянному возмущению.

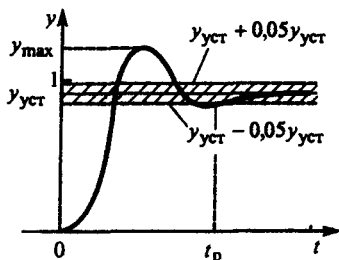


Рис. 3.25. Прямые показатели качества канала воспроизведения задания

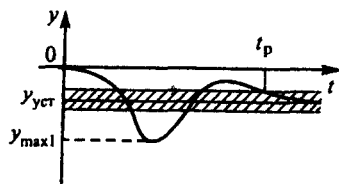


Рис. 3.26. Прямые показатели качества канала подавления возмущения

В отличие от оценок свободных движений, определяемых только расположением корней характеристического полинома, здесь в общем случае оцениваются расположения полюсов и нулей передаточной функции по тому или иному каналу. Как показано в § 2.4, приближение нулей передаточной функции к полюсам, т. е. образование диполей (неполнота характеристик), уменьшает переходные составляющие процессов при любых воздействиях.

Переходные составляющие процессов при нулевых предначальных условиях

$$y_{\text{пер}}(t) = y(t) - y_{\text{уст}}$$

являются собственными (сопровождающими) составляющими вынужденных движений. Они могут рассматриваться как решение эквивалентного однородного дифференциального уравнения при посленаачальных условиях, определяемых и правой частью неоднородного дифференциального уравнения. Для количественной характеристики этих процессов также применяют интегральные оценки, рассмотренные в § 3.11.

Косвенные частотные показатели качества передач вводятся на базе амплитудно-частотных или вещественных частотных характеристик канала. Традиционно при анализе и синтезе следящих систем эти характеристики рассматриваются по каналу воспроизведения. На рис. 3.27 изображена такая амплитудно-частотная характеристика. Показателями качества для амплитудно-частотной характеристики являются показатель колебательности M и частота резонанса ω_p . Значение $R_g(0)$ равно установившемуся значению $y_{уст}$ при постоянном единичном сигнале. Как видно из рис. 3.27, выход системы ковариантен до ϵ с задающими сигналами низких частот.

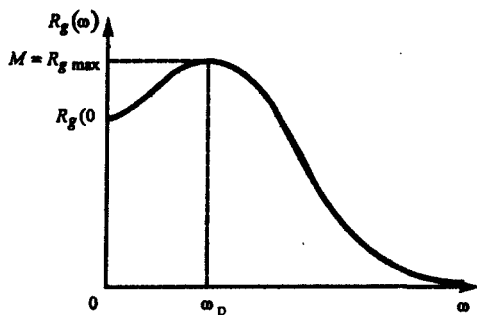


Рис. 3.27. Частотные показатели качества канала воспроизведения задания

Амплитудно-частотная характеристика, отражающая установившуюся реакцию системы на гармонические сигналы всех частот, позволяет судить и о динамических свойствах минимально-фазовой системы управления [75].

3.13. УПРАВЛЯЕМОСТЬ И НАБЛЮДАЕМОСТЬ СИСТЕМ

3.13.1. Понятия управляемости и наблюдаемости

В случае многомерных систем (объектов), представленных в форме пространства состояний (2.18), понятие полноты, рассмотренное в § 2.4 применительно к одномерным системам, расширяется. Вводятся понятия *управляемости* и *наблюдаемости*.

На содержательном уровне управляемость означает принципиальную возможность перевода системы или объекта из одного состояния в другое за конечное время, а наблюдаемость — возможность определе-

ния состояния системы или объекта за конечное время по информации о входах и выходах. Эти свойства необходимо учитывать при построении большинства систем управления.

Для установления свойств управляемости и наблюдаемости рассматривается система уравнений вида

$$\begin{aligned}\frac{dv}{dt} &= Av + Bf, \\ y &= Cv.\end{aligned}\tag{3.21}$$

Поскольку матрица D не оказывает влияния на свойства управляемости и наблюдаемости, ее можно не учитывать в описании, т. е. считать нулевой. Вектор f в (3.21) представляет собой вектор управляющих воздействий. Векторно-матричная диаграмма системы S , соответствующая описанию (3.21), приведена на рис. 3.28.

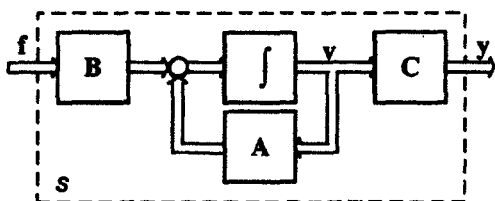


Рис. 3.28. Векторно-матричная диаграмма

Рассматривают управляемость как динамических объектов без регуляторов, так и систем, содержащих замкнутые контуры управления. Существуют различные определения понятия управляемости, различающиеся как условиями перевода системы, так и ограничениями, накладываемыми на управляющие воздействия. Наибольшее распространение получили определения понятий управляемости и наблюдаемости, предложенные Р. Калманом (R. Kalman).

Система (объект) называется *полностью* или *вполне управляемой*, если существует вектор воздействий $f(t)$, который за конечное время $\Delta t = t_* - t_0$ переводит систему (объект) из любого исходного $v(t_0)$ в любое заданное $v(t_*)$ состояние. Обычно вектор $f(t)$ принадлежит к классу кусочно-непрерывных функций.

Для перевода системы за бесконечное время достаточно, чтобы конечное состояние $v(t_*)$ было асимптотически устойчивым «в целом», т. е. для всего пространства состояний (см. § 12.1). Перевод за конечное время требует, кроме того, выполнения дополнительных условий управляемости.

Более сильной формой управляемости является *нормализуемость* (*нормальность*). Говорят, что система является нормализуемой, если каждая координата вектора управляющих воздействий $\mathbf{f}(t)$ в отдельности обеспечивает управляемость.

Система (объект) называется *полностью* или *вполне наблюдаемой*, если по данным измерений векторов $\mathbf{y}(t)$ и $\mathbf{f}(t)$ можно за конечное время $\Delta t = t_* - t$ однозначно определить состояние системы $\mathbf{v}(t)$. Другими словами, система полностью наблюдаема, если существует момент времени t_* , $t < t_* < \infty$ такой, что по данным измерений $\mathbf{y}(\tau)$, $\mathbf{f}(\tau)$, $\tau \in [t, t_*]$ можно определить состояние $\mathbf{v}(t)$. Полная наблюдаемость системы означает возможность определения состояния $\mathbf{v}(t)$ по будущим значениям векторов $\mathbf{y}(\tau)$, $\mathbf{f}(\tau)$, что практически не осуществимо. Поэтому в задачах управления более важным обстоятельством является установление свойства полной восстанавливаемости вектора $\mathbf{v}(t)$ по прошлым значениям векторов $\mathbf{y}(\tau)$, $\mathbf{f}(\tau)$.

Система (объект) называется *полностью* или *вполне восстанавливаемой*, если существует такой момент времени t_0 , $t_0 < t < \infty$, что по данным измерений $\mathbf{y}(\tau)$, $\mathbf{f}(\tau)$, $\tau \in [t_0, t]$ можно определить состояние $\mathbf{v}(t)$.

Для линейных стационарных систем из свойства полной наблюдаемости следует свойство полной восстанавливаемости и наоборот.

Однако не любая система является полностью управляемой или полностью наблюдаемой. Предположим, что в результате преобразования система S , отвечающая (3.21), в новом базисе переменных состояния $\tilde{\mathbf{v}}$ может быть представлена в виде

$$\begin{aligned}\frac{d\tilde{\mathbf{v}}_1}{dt} &= \mathbf{A}_{11}\tilde{\mathbf{v}}_1 + \mathbf{A}_{12}\tilde{\mathbf{v}}_2 + \mathbf{B}_1\mathbf{f}; \\ \frac{d\tilde{\mathbf{v}}_2}{dt} &= \mathbf{0}\tilde{\mathbf{v}}_1 + \mathbf{A}_{22}\tilde{\mathbf{v}}_2; \\ \mathbf{y} &= \mathbf{C}_1\tilde{\mathbf{v}}_1 + \mathbf{C}_2\tilde{\mathbf{v}}_2.\end{aligned}\tag{3.22}$$

Приведенная запись означает декомпозицию системы S (3.21) на подсистемы S_1 , S_2 (рис. 3.29, а). В преобразованной системе уравнений (3.22) принято $\tilde{\mathbf{v}} = (\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2)^T$, $\tilde{\mathbf{v}}_1$, $\tilde{\mathbf{v}}_2$ — подвекторы вектора $\tilde{\mathbf{v}}$ подсистем S_1 , S_2 соответственно; $\mathbf{A}_{12}$ — матрица направленных связей на подсистему S_1 со стороны подсистемы S_2 . Особенность уравнения состояния подсистемы S_2 состоит в том, что в него не входит ни воздействие $\mathbf{f}$, ни переменные состояния, относящиеся к подвектору $\tilde{\mathbf{v}}_1$ подсистемы S_1 , в уравнение состояния которой входит $\mathbf{f}$. Другими словами, управляющее воздействие $\mathbf{f}$ не оказывает влияния на переменные подвектора $\tilde{\mathbf{v}}_2$ подси-

системы S_2 ни непосредственно, ни опосредованно через переменные подсистемы S_1 . Это означает, что состояние $\bar{V}_2$ не управляется воздействием f . Следовательно, система не полностью управляема по вектору состояния.

Для управления системой по принципу обратной связи нередко требуется располагать информацией о текущем состоянии системы, т. е. о значениях переменных состояния вектора $v(t)$ в каждый момент времени. Однако не все переменные состояния имеют физический аналог в реальной системе. Некоторые из них представляют абстрактные величины, а потому не могут быть измерены. Измеряемыми являются физические переменные вектора выхода $y(t)$, через которые должны однозначно выражаться все переменные вектора состояния $v(t)$. В линейных системах связь между v и y линейная. В общем случае вектор y может зависеть не только от вектора v , но и от вектора управляющих воздействий f .

Пусть используется такое преобразование, что часть переменных состояния $\bar{V}$ системы S не участвует непосредственно или опосредованно в формировании переменной выхода y , т. е.

$$\begin{aligned}\frac{d\bar{V}_1}{dt} &= A_{11}\bar{V}_1 + A_{12}\bar{V}_2 + B_1 f; \\ \frac{d\bar{V}_2}{dt} &= 0\bar{V}_1 + A_{22}\bar{V}_2 + B_2 f; \\ y &= 0\bar{V}_1 + C_2\bar{V}_2.\end{aligned}\tag{3.23}$$

Преобразованная система (3.23) (рис. 3.29, б) не полностью наблюдаема по вектору состояния $\bar{V}$, поскольку выход системы y не содержит переменных вектора $\bar{V}_1$. На состояние вектора $\bar{V}_2$ подсистемы S_2 состояние вектора $\bar{V}_1$ подсистемы S_1 ни непосредственно, ни опосредованно не влияет.

На рис. 3.29, в и г приведены диаграммы полностью ненаблюдаемой и полностью неуправляемой систем.

В общем случае система может состоять из четырех частей:

- 1) полностью неуправляемой и полностью ненаблюдаемой;
- 2) полностью управляемой, но не полностью наблюдаемой;
- 3) полностью наблюдаемой, но не полностью управляемой;
- 4) полностью управляемой и полностью наблюдаемой.

Если система полностью и управляема, и наблюдаема, то ее называют *полной*.

В одномерном случае передаточная функция системы $W(s)$ является невырожденной тогда и только тогда, когда одновременно выполняются условия полной управляемости и полной наблюдаемости.

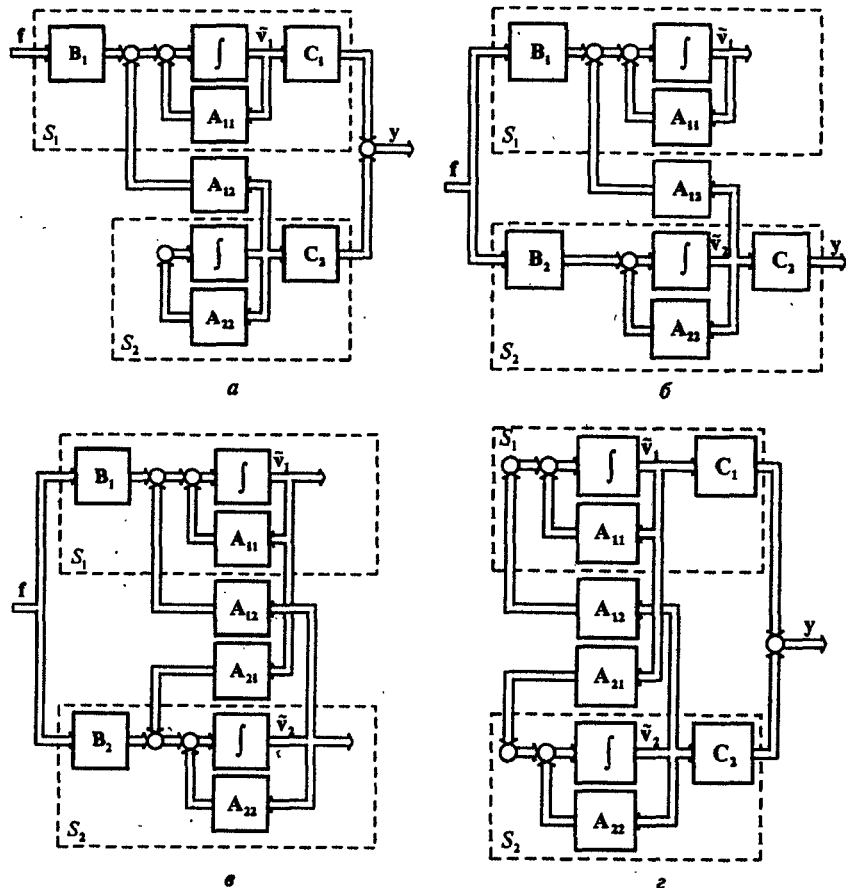


Рис. 3.29. Диаграммы систем:

a — не полностью управляемой; *б* — не полностью наблюдаемой;
в — полностью ненаблюдаемой; *г* — полностью неуправляемой

3.13.2. Критерии управляемости и наблюдаемости

Для установления свойств управляемости и наблюдаемости системы (объекта), представленной в форме пространства состояний, используются *алгебраические критерии* Р. Кадмана. С этой целью составляются матрица управляемости

$$U = (B \quad A \quad B \quad A^2 B \dots A^{n-1} B)$$

размера $n \times nr$ (для скалярного входа f матрица B — вектор-столбец, при этом U — квадратная матрица размера $n \times n$) и матрица наблюдаемости

$$V = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

размера $kn \times n$ (для скалярного выхода у матрица C — вектор-строка, при этом V — квадратная матрица размера $n \times n$).

Критерий управляемости. Система полностью управляема тогда и только тогда, когда матрица управляемости имеет полный ранг, т. е.

$$\text{rank} U = n. \quad (3.24)$$

Для скалярного входа f критерий приобретает вид

$$\det U \neq 0.$$

В соответствии с критерием (3.24) свойство управляемости целиком определяется парой матриц (A, B) , поэтому часто говорят об управляемости пары (A, B) .

Пара (A, B) , в которой A и B — матрицы размеров $n \times n$ и $n \times p$ соответственно называется полностью управляемой, если система (3.21) полностью управляема.

Для установления свойства нормализуемости («сильной» управляемости) составляются матрицы управляемости:

$$U_i = (B_i \quad AB_i \quad A^2B_i \dots A^{n-1}B_i); \quad i = 1, \dots, p,$$

где B_i — столбцы матрицы B .

Критерий нормализуемости. Система полностью нормализуема тогда и только тогда, когда для всех $i = 1, \dots, p$ выполняются условия

$$\text{rank} U_i = n.$$

Очевидно, система управляема, если выполняется критерий нормализуемости, однако обратное утверждение неверно. При скалярном управлении оба критерия совпадают.

Критерий наблюдаемости. Система полностью наблюдаема тогда и только тогда, когда матрица наблюдаемости имеет полный ранг, т. е.

$$\text{rank} V = n. \quad (3.25)$$

Для скалярного выхода у

$$\det V \neq 0.$$

В соответствии с критерием (3.25) свойство наблюдаемости целиком определяется парой матриц (A, C) , поэтому часто говорят о наблюдаемости пары (A, C) .

Пара (A, C), в которой A и C — матрицы размеров $n \times n$ и $k \times n$ соответственно, называется полностью наблюдаемой, если система (3.21) полностью наблюдаема.

Наблюдаемость систем часто устанавливают по транспонированной матрице

$$V^T = (C^T \ A^T C^T \ (A^T)^2 C^T \ \dots \ (A^T)^{n-1} C^T)$$

размера $n \times nk$, которая имеет такой же ранг, что и матрица V.

Рассмотрим две системы S и $\tilde{S}$, для которых переменные состояния связаны неособым линейным преобразованием $v = Tv$ так, что

$$\tilde{A} = T^{-1}AT, \quad \tilde{B} = T^{-1}B, \quad \tilde{C} = CT.$$

Тогда для преобразованной системы $\tilde{S}$ матрица управляемости

$$\begin{aligned} \tilde{U} &= (\tilde{B} \ \tilde{A}\tilde{B} \ \dots \ \tilde{A}^{n-1}\tilde{B}) = \\ &= (T^{-1}B \ T^{-1}ATT^{-1}B \ \dots \ T^{-1}ATT^{-1}AT \ \dots \ T^{-1}ATT^{-1}B) = \\ &= (B \ AB \ \dots \ A^{n-1}B) = T^{-1}U; \end{aligned}$$

матрица наблюдаемости

$$\tilde{V} = \begin{pmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \vdots \\ \tilde{C}\tilde{A}^{n-1} \end{pmatrix} = \begin{pmatrix} CT \\ CT T^{-1}AT \\ \vdots \\ CT T^{-1}AT \dots T^{-1}AT \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} T = VT.$$

Ранги матриц $\tilde{U}$ и U, $\tilde{V}$ и V одинаковые.

Таким образом, при неособом линейном преобразовании системы ранги матриц U и V не изменяются, а следовательно, остаются неизменными свойства управляемости и наблюдаемости. Другими словами, свойства управляемости и наблюдаемости не зависят от выбора переменных состояний.

В том случае, если для одномерных систем S и $\tilde{S}$ известны их матрицы A и B, $\tilde{A}$ и $\tilde{B}$, передаточные функции обеих систем совпадают, то неособая матрица преобразования T определяется выражением

$$T = U\tilde{U}^{-1}.$$

Пример 3. Система управления описывается передаточной функцией

$$W(s) = \frac{s+a}{(s+b)(s+c)}. \quad (3.26)$$

При $b \rightarrow a$ или $c \rightarrow a$ образуется диполь передаточной функции. Можно показать, что в зависимости от выбора базиса переменных в пространстве состояний, система либо не полностью управляема, либо не полностью наблюдаема.

Матрица состояний в форме Фробениуса, составленная по $W(s)$, имеет вид

$$A = \begin{pmatrix} 0 & 1 \\ -bc & -(b+c) \end{pmatrix}.$$

В каноническом управляемом базисе (см. п. 2.12.1):

$$B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \quad C = (a \ 1);$$

$$AB = \begin{pmatrix} 0 & 1 \\ -bc & -(b+c) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -(b+c) \end{pmatrix};$$

$$CA = (a \ 1) \begin{pmatrix} 0 & 1 \\ -bc & -(b+c) \end{pmatrix} = (-bc \ a - (b+c)).$$

В соответствии с критериями управляемости и наблюдаемости

$$\det U = \det \begin{pmatrix} 0 & 1 \\ 1 & -(b+c) \end{pmatrix} = -1,$$

$$\det V = \det \begin{pmatrix} a & 1 \\ -bc & a - (b+c) \end{pmatrix} = (a-b)(a-c).$$

При $b \rightarrow a$ или $c \rightarrow a$ определить $\det V = 0$, т. е. система управляема, но не полностью наблюдаема.

В каноническом наблюдаемом базисе:

$$B = \begin{pmatrix} 1 \\ a - (b+c) \end{pmatrix}; \quad C = (1 \ 0);$$

$$AB = \begin{pmatrix} 0 & 1 \\ -bc & -(b+c) \end{pmatrix} \begin{pmatrix} 1 \\ a - (b+c) \end{pmatrix} = \begin{pmatrix} a - (b+c) \\ -bc - (b+c)(a-b-c) \end{pmatrix};$$

$$CA = (1 \ 0) \begin{pmatrix} 0 & 1 \\ -bc & -(b+c) \end{pmatrix} = (0 \ 1).$$

В соответствии с критериями управляемости и наблюдаемости

$$\det U = \det \begin{pmatrix} 1 & a - (b+c) \\ a - (b+c) & -bc - (b+c)(a-b-c) \end{pmatrix} = (a-b)(c-a);$$

$$\det V = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

При $b \rightarrow a$ или $c \rightarrow a$ определитель $\det U = 0$, т. е. система наблюдаема, но не полностью управляема.

Следовательно, если система по рассматриваемым входу и выходу неполная (передаточная функция содержит диполи), то она либо не полностью управляема, либо не полностью наблюдаема.

3.13.3. Принцип дуальности

Пусть рассматриваются две системы

$$\begin{aligned}\frac{dv}{dt} &= Av + Bf; \\ y &= Cv\end{aligned}\tag{3.27}$$

и

$$\begin{aligned}\frac{dw}{dt} &= A^T w + C^T g; \\ z &= B^T w.\end{aligned}\tag{3.28}$$

Для таких систем, которые называются *двойственными* (сопряженными), Р. Калманом был установлен *принцип дуальности*:

условие полной управляемости системы (3.27), т. е.

$$\text{rank}(B \ AB \ A^2 B \dots A^{n-1} B) = n$$

означает полную наблюдаемость двойственной системы (3.28), т. е.

$$\text{rank} \begin{pmatrix} B^T \\ B^T A^T \\ B^T (A^T)^2 \\ \vdots \\ B^T (A^T)^{n-1} \end{pmatrix} = n;$$

условие полной наблюдаемости системы (3.27), т. е.

$$\text{rank} \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

означает полную управляемость двойственной системы (3.28), т. е.

$$\text{rank} \begin{pmatrix} C^T & A^T C^T & (A^T)^2 C^T & \dots & (A^T)^{n-1} C^T \end{pmatrix} = n.$$

Другими словами, система управляема, если наблюдаема двойственная ей система, и наоборот. Следовательно, для того, чтобы установить

наблюдаемость системы (3.27), достаточно установить управляемость двойственной ей системы (3.28), что облегчает разработку алгоритмов определения управляемости и наблюдаемости.

Пример 4. Установим выполнение принципа дуальности по отношению к системе, представленной передаточной функцией (3.26). Располагая описанием исходной системы в пространстве состояний (см. пример 3), получим описание двойственной системы в соответствии с (3.28).

Матрица состояний

$$A^T = \begin{pmatrix} 0 & -bc \\ 1 & -(b+c) \end{pmatrix}.$$

Запишем двойственную систему, например, в каноническом управляемом базисе:

$$B^T = (0 \ 1); \quad C^T = \begin{pmatrix} a \\ 1 \end{pmatrix};$$

$$A^T C^T = \begin{pmatrix} 0 & -bc \\ 1 & -(b+c) \end{pmatrix} \begin{pmatrix} a \\ 1 \end{pmatrix} = \begin{pmatrix} -bc \\ a - (b+c) \end{pmatrix};$$

$$B^T A^T = (0 \ 1) \begin{pmatrix} 0 & -bc \\ 1 & -(b+c) \end{pmatrix} = (1 - (b+c)).$$

Если исходная система полностью управляема, то двойственная ей система полностью наблюдаема, поскольку

$$\det \begin{pmatrix} 0 & 1 \\ 1 & -(b+c) \end{pmatrix} = -1.$$

Если исходная система полностью наблюдаема, то двойственная ей система полностью управляема, поскольку

$$\det \begin{pmatrix} a & -bc \\ 1 & a - (b+c) \end{pmatrix} = (a-b)(a-c) \neq 0$$

при $a \neq b, a \neq c$.

Глава 4

СИНТЕЗ СИСТЕМ УПРАВЛЕНИЯ

4.1. О СИНТЕЗЕ СИСТЕМ УПРАВЛЕНИЯ

Целью синтеза является построение математической модели системы управления, удовлетворяющей требованиям к поведению: ковари-

антности с заданием; инвариантности к возмущениям; устойчивости и грубости (робастности).

Пусть имеется математическая модель объекта управления со связями со средой (рис. 4.1). В задачах синтеза алгоритмов управления к объекту ОУ или неизменяемой части относят исполнительные механизмы ИМ и измерительные элементы ИЭ, как это показано на рис. 4.1. При этом входом $u(t)$ и выходом $y(t)$ расширенного объекта оказываются маломощные сигналы — носители информации.

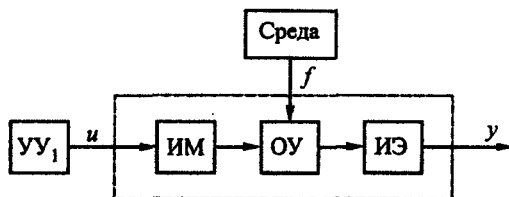


Рис. 4.1. Расширенный объект управления

Средствами решения задачи синтеза в указанной постановке являются: выбор структур систем (т. е. элементов и топологии причинно-следственных связей между ними), структур операторов элементов (в частности, алгоритмов управляющих устройств) и значений их параметров (например, настроек регуляторов).

Удовлетворению требований к поведению систем обычно препятствуют динамические свойства объектов управления и других элементов неизменяемой части, недоступность полной априорной информации о свойствах элементов системы и среды, невозможность получения всей текущей информации о состоянии объекта и возмущениях, ограничения на переменные системы и управляющие воздействия.

Более общей по сравнению с синтезом является задача *проектирования* систем управления. Несмотря на то что требования к поведению систем управления являются доминирующими, при проектировании необходимо учитывать и другие требования и ограничения, содержащиеся в технических заданиях: надежность систем, их приемлемую стоимость, требования энергетического характера, ограничения, связанные с типом сигналов, массой и габаритами систем, компоновкой элементов и трассировкой связей и т. д. Для расчетов систем с учетом таких требований привлекаются модели и методы, отличные от рассматриваемых в традиционных курсах теории управления.

Для решения задач анализа — проверки соответствия поведения системы требованиям — необходимо иметь описание системы, среды

и требований. Если система не удовлетворяет требованиям, то принимается решение о необходимости синтеза.

В постановках задач синтеза задается *множество систем* M , на котором проводится *выбор*. Задача синтеза нетривиальна только в случае, когда множество M содержит более одного элемента, иначе говоря, имеется исходная неопределенность. Выделим следующие постановки задачи.

• Во-первых, элементы множества M могут различаться параметрами. При этом множество M является моделью второго ранга неопределенности

$$M = M(2) = \{M(3)\},$$

т. е. множеством полностью определенных систем третьего ранга. На рис. 2.22, *а* представлена ситуация, когда системы различаются параметрами одного из звеньев, например настройками регулятора. В результате *параметрического синтеза* находятся их значения.

• Во-вторых, элементы исходного множества M могут различаться структурами операторов звеньев. При этом множество M является моделью первого ранга неопределенности

$$M = M(1) = \{M(2)\},$$

т. е. множеством моделей второго ранга. В результате *структурного синтеза* выбирается подмножество структур операторов (в частном случае — единственная структура), после чего имеет место рассмотренная выше задача параметрического синтеза. На рис. 2.22, *б* иллюстрируется ситуация, когда оператор одного из звеньев системы может иметь различные структуры. В частном случае этим звеном может быть регулятор, а в результате структурного синтеза необходимо найти его тип.

• В-третьих, множество M может быть моделью нулевого ранга неопределенности

$$M = M(0) = \{M(1)\},$$

т. е. представляет собой множество систем с различной топологией. В результате *топологического синтеза* выбирается подмножество топологий (в частности, единственная топология), после чего решается задача структурного синтеза. На рис. 2.22, *в* иллюстрируется ситуация, когда системы различаются местом включения одного звена. В результате синтеза необходимо найти наилучшее место включения регулятора или компенсатора, т.е. определить точки измерения и оказания на объект управляющих воздействий.

Синтез представляет собой повышение ранга моделей R , т.е. уменьшение неопределенности или разнообразия за счет привлечения инфор-

мации о пожеланиях проектировщика. Требования удобно рассматривать как множество систем M^d , поведение которых удовлетворительно. Описание этого множества, в отличие от «структурно-операторного» описания систем, строится на «поведенческом» языке, характеризуя качественно и количественно свойства устойчивости, инвариантности и робастности.

4.2. ЗАДАЧИ СИНТЕЗА СИСТЕМ УПРАВЛЕНИЯ

Рассмотрим на содержательном уровне основные задачи синтеза систем управления.

Синтез управляющих воздействий. Приложенное к объекту допустимое управляющее воздействие $u(t) \in \mathcal{U}$ должно обеспечить в некотором смысле наилучшее поведение объекта. Задача синтеза математически ставится как поиск функции времени $u^*(t)$, доставляющей минимум (максимум) некоторому функционалу

$$J(y(t), u(t), f(t)) \rightarrow \min(\max)$$

с учетом динамических свойств объекта, ограничений на его переменные состояния, а также тех возмущений $f(t)$, о которых имеется полная априорная информация. Решение сводят к поиску условных экстремалей функционалов, для чего привлекаются методы классического вариационного исчисления, динамического программирования и принцип максимума. Эти задачи и методы их решения будут изложены далее.

Управляющее воздействие $u^*(t)$ генерируется управляющим устройством УУ₁ (см. рис. 4.1) и обеспечивает оптимальную траекторию движения объекта $y^*(t)$.

Во многих технических объектах оптимальное управление постоянно $u^* = \text{const}$ и обеспечивает оптимальный режим $y^* = \text{const}$, определяемый из требований технологии.

Синтез компенсаторов возмущений. Если на объект действуют возмущения $f(t)$, которые не учтены при синтезе оптимального управления $u^*(t)$, то поведение объекта будет отличаться от оптимального. В случае недопустимых отклонений соответствующей траектории (режима) необходимо принять меры по ослаблению влияния возмущений.

Пусть возмущение измеряется непосредственно (рис. 4.2). Задачей синтеза является определение алгоритма управляющего устройства УУ₂, в котором происходит обработка текущей информации о возмущении и формирование воздействия $u_2(t)$ на объект. Часть управляющего устройства, формирующую компенсирующее воздействие, называют также компенсатором (К).

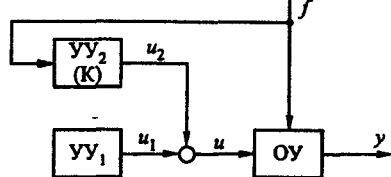


Рис. 4.2. Компенсация
возмущения

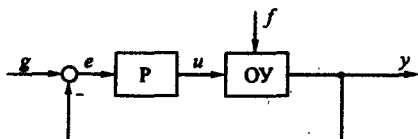


Рис. 4.3. Система с внутренним
компенсатором

Образование канала компенсации, в принципе, может обеспечить абсолютную инвариантность управляемой переменной к непосредственно измеряемому возмущению.

Синтез систем управления из условия подавления непосредственно неизмеряемых возмущений. Если к объекту приложены постоянно действующие некомпенсированные возмущения $f(t)$, то единственным средством ослабления их влияния на управляемую переменную $y(t)$ является создание контуров обратной связи с достаточно большим усилением на частотах возмущений. Необходимое усиление контура обеспечивается введением *внутренних компенсаторов* возмущений (рис. 4.3).

Селективная абсолютная инвариантность достигается, если внутренний компенсатор «моделирует среду» — его передаточная функция имеет полюсы, равные полюсам изображения возмущения с учетом их кратности (см. п. 3.7.3). Из-за бесконечного усиления контура на комплексных частотах возмущения установившаяся реакция системы равна нулю. Селективная инвариантность до ϵ обеспечивается, если на частотах возмущений усиления контура достаточно велики.

Пусть измеряются непосредственно выходная переменная объекта и основное возмущение. Тогда создают *комбинированные* системы управления, реализующие оба принципа управления — по замкнутому и разомкнутому циклам. Канал компенсации обеспечивает инвариантность к основному возмущению (например, ток нагрузки электрического генератора, момент сопротивления на валу двигателя), а обратная связь ослабляет влияние непосредственно неизмеряемых возмущений, а также вариаций операторов.

Компенсация главного возмущения позволяет снизить требования к астатизму системы и усилению контура обратной связи. Это, в свою очередь, облегчает решение задачи стабилизации системы и удовлетворения требований к переходным процессам.

Синтез следящих систем управления. В так называемых следящих системах управляемая переменная $y(t)$ должна воспроизводить заранее неизвестное воздействие $g(t)$, т. е. должна быть ковариантной с ним. Текущая информация обычно может быть получена только об ошибке слежения

$$e(t) = g(t) - y(t).$$

В этом случае строится система с обратной связью (см. рис. 4.3), а задачей синтеза является определение алгоритма регулятора

$$u(t) = R(e(t)),$$

обеспечивающего воспроизведение задающего воздействия с требуемой точностью. Переменная ошибки $e(t)$ должна быть инвариантной к заданию $g(t)$ и возмущениям $f(t)$.

В следящих системах без непосредственного измерения задающего воздействия нельзя реализовать абсолютную инвариантность ошибки к заданию. Селективная абсолютная инвариантность достигается, если ввести в контур внутренний компенсатор воздействия (см. рис. 4.3), полюсы передаточной функции которого равны полюсам изображения воздействия $g(t)$.

Легко видеть, что в обеих системах — подавления воздействий и воспроизведения задания — переменная ошибки должна быть инвариантной к воздействиям. Разница между задачами синтеза этих систем заключается в том, что основные требования в первом случае относятся к точности подавления возмущения, а во втором — к точности воспроизведения задания.

Коррекция систем управления. Создание контуров обратной связи и повышение их усиления, введение в контуры внутренних компенсаторов, передаточные функции которых имеют полюсы, равные полюсам изображений воздействий, обычно приводит к тому, что замкнутая система с удовлетворительными установившимися движениями будет иметь плохие переходные процессы или даже окажется неустойчивой. Это часто имеет место, когда спектры воздействий близки к спектру объекта, т. е. усиления контура повышаются на частотах объекта, вызывая существенное перемещение полюсов передаточной функции объекта.

Если средства, обеспечивающие инвариантность переменной ошибки к непосредственно неизмеряемым воздействиям, приводят к неустойчивости замкнутой системы, то необходимо разрешить это противоречие между качеством установившихся и переходных процессов. Такая задача синтеза называется *коррекцией*.

Синтез систем с обратной связью в общем случае является весьма сложной задачей. Ее сложность обусловлена разнообразием требований: одновременно необходимо обеспечить инвариантность к возмущениям, ковариантность с заданием, устойчивость движений, а также, по возможности, малую чувствительность перечисленных свойств. Эти требования обычно оказываются противоречивыми, что превращает процедуру синтеза регуляторов в последовательность принятия компромиссных решений.

Частично упростить задачу синтеза возможно в случае линейных моделей, когда установившиеся и переходные составляющие движений в какой-то мере можно формировать раздельно. Декомпозиция возможна при условии, когда собственные движения системы более быстрые по сравнению с воздействиями среды. Это означает, что модули полюсов передаточной функции системы должны быть больше модулей полюсов воздействий.

Таким образом, процедура синтеза систем подавления возмущений и воспроизведения задания складывается из двух основных этапов:

- синтеза компенсатора по требованиям к установившимся процессам;
- коррекции по требованиям к переходным процессам.

При этом управляющее устройство представляется как совокупность компенсатора воздействия, обеспечивающего установившуюся точность, и звена коррекции, обеспечивающего устойчивость и требуемое качество переходных процессов.

Синтез регуляторов для неустойчивых объектов. Оптимальные траектории $y^*(t)$, в частном случае — оптимальный режим $y^* = \text{const}$ — могут быть неустойчивыми или вариации движений могут затухать недостаточно быстро. Тогда ставится задача стабилизации неустойчивого режима и обеспечения требуемого характера переходных процессов.

Для изменения характера собственных движений необходимо создать систему с обратной связью (рис. 4.4, а), т.е. реализовать принцип управления по замкнутому циклу. Управляющее устройство УУ₃, обеспечивающее устойчивость и качество процессов в окрестности оптимального режима, называют также регулятором (Р).

Задачей синтеза в этом случае является определение алгоритма регулирования

$$\delta u(t) = R(\delta y(t)),$$

а именно, его типа (структуры) и настроек (параметров).

Вообще говоря, стабилизирующая обратная связь может включаться и иначе (рис. 4.4, б): измеряется некоторая внутренняя переменная $x(t)$ и

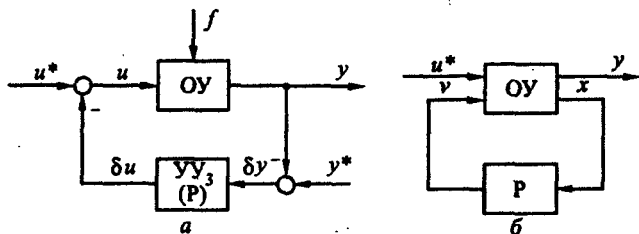


Рис. 4.4. Стабилизация неустойчивого режима

на объект оказывается воздействие по дополнительному входу $v(t)$. При наличии нескольких мест возможного включения регулятора возникает задача топологического синтеза — выбора наилучшего места.

Во многих практически важных случаях задача стабилизации решается по математическим моделям, линеаризованным для малых отклонений от рассматриваемого режима.

Расчет настроек типовых регуляторов. Большинство локальных систем управления промышленной автоматики имеют типовую одноконтурную структуру (см. рис. 4.3) с управляющими устройствами, реализующими типовые алгоритмы (см. п. 1.5.2). Это объясняется универсальностью обратной связи и тем, что динамические свойства большинства объектов также принадлежат к нескольким типам. Таким образом, системы имеют определенную топологию и структуры операторов, но остаются неопределенными параметры управляющих устройств, которые в системах стабилизации и следящих системах обычно называются регуляторами.

Параметры настроек типовых регуляторов рассчитываются из условий устойчивости, максимизации точности поддержания установившихся режимов, минимизации динамических ошибок.

Предложены рекомендации по выбору настроек типовых регуляторов и разработаны графо-аналитические методики их расчета [12, 68]. В настоящее время для оптимизации систем с типовыми регуляторами привлекаются ЭВМ [4, 30].

Синтез систем управления в условиях неполной определенности моделей. Даже при точной реализации алгоритма управления, синтезированного на базе полностью определенной модели, реальная система, вообще говоря, будет иметь другое поведение, так как реальная динамика объекта отличается от модельной. Поэтому любой метод синтеза имеет смысл только в том случае, если он по меньшей мере гарантирует, что малые вариации характеристик элементов не вызовут боль-

ших изменений поведения систем. Синтезированная система должна быть грубой — это необходимое условие применимости методов синтеза. Практически же требуется робастность основных свойств — система управления должна быть работоспособной при конечных изменениях характеристик элементов.

Передачи замкнутых систем малочувствительны к вариациям характеристик некоторых элементов на частотах, где усиления контуров велики (см. п. 3.9.3). Наличие контура является необходимым структурным (топологическим) условием стабилизации неустойчивых объектов, ослабления сигнальных и операторных возмущений. В этом состоит универсальность действия обратной связи.

При синтезе систем частотными методами можно контролировать диапазоны частот, на которых вариации характеристик элементов оказывают незначительное влияние на характеристики систем. Для количественной оценки влияния вариаций элементов на условия инвариантности систем можно использовать функции чувствительности (см. § 3.9).

Если вариации элементов велики и/или усиления контуров на требуемых частотах не могут быть повышены, то необходимы активные средства ослабления влияния вариаций элементов на свойства систем. В этих случаях строят адаптивные системы управления, в которых за счет обработки текущей информации о динамических свойствах объектов происходит перестройка алгоритма управляющего устройства.

Задачи синтеза адаптивных систем решаются вне рамок линейной теории управления.

4.3. СТАБИЛИЗАЦИЯ НЕУСТОЙЧИВЫХ ОБЪЕКТОВ

Пусть анализ линейной модели, описывающей движения объекта в окрестности выбранного режима работы, показал, что режим неустойчив. Математически этот факт выражается в том, что характеристический полином дифференциального уравнения имеет корни с неотрицательными действительными частями. Возникает задача стабилизации неустойчивого объекта. В других случаях объект может быть устойчивым, но его собственные движения не удовлетворяют требованиям (например, движения сильно колебательные или/и затухают слишком медленно).

Необходимым топологическим условием изменения расположения корней характеристического полинома является образование контура, содержащего объект управления (см. п. 3.3.2). Кроме того, передаточная функция объекта по выбранному каналу вход-выход не должна иметь

неустойчивых диполей. В противном случае никакая обратная связь не сможет переместить корни неполной части.

В зависимости от формы представления модели объекта и требований к собственным движениям системы могут быть применены различные методы синтеза.

4.3.1. Размещение корней характеристического полинома.

Операторный метод

Допустим, что требования к системе представлены в форме желаемого множества корней характеристического полинома. Необходимо найти алгоритм регулятора, размещающего корни в назначенных местах комплексной плоскости. Корни характеристического полинома $\{p_i\}$ — это полюсы передаточной функции системы; по этой причине иногда говорят о задаче размещения полюсов. Поскольку корням p_i соответствуют составляющие собственных движений $\exp\{p_i t\}$ — так называемые *моды* — задачу размещения корней иногда называют управлением модами, или *модальным управлением* [5].

Запишем дифференциальное уравнение объекта в операторной форме

$$A_0(p)y(t) = B_0(p)u(t). \quad (4.1)$$

Положим, что степень n_0 полинома A_0 выше степени m_0 полинома B_0 . Кроме того, допустим, что полиномы A_0 и B_0 взаимно просты, т. е. описание вход-выход объекта (4.1) является *полным* (см. § 2.4). Без потери общности примем, что коэффициент при старшей степени полинома A_0 равен единице.

Искомое дифференциальное уравнение стабилизирующей отрицательной обратной связи (регулятора) также запишем в общем виде в операторной форме

$$A_R(p)u(t) = -B_R(p)y(t). \quad (4.2)$$

Однородное дифференциальное уравнение автономной замкнутой системы получим, если исключим переменную $u(t)$ из уравнений (4.1) и (4.2):

$$A_3(p)y(t) \equiv [A_0(p)A_R(p) + B_0(p)B_R(p)]y(t) = 0. \quad (4.3)$$

Потребуем тождества характеристического полинома $A_3(p)$ желаемому полиному

$$A^*(p) = \prod_i (p - p_i^*), \quad (4.4)$$

построенному по заданным корням p_i^* ; $i = 1, \dots, n$:

$$A_0(p)A_R(p) + B_0(p)B_R(p) \equiv A^*(p). \quad (4.5)$$

Из тождества (4.5) необходимо найти операторные полиномы регулятора $A_R(p)$ и $B_R(p)$. Это значит, что следует искать структуру регулятора — степени n_R и m_R , а также параметры регулятора — коэффициенты полиномов:

$$A_R(p) = p^{n_R} + \dots + a_{1R}p + a_{0R};$$

$$B_R(p) = b_{m_R}p^{m_R} + \dots + b_{1R}p + b_{0R}.$$

Полиномиальные уравнения вида (4.5) называют *диофантовыми*, так как полиномы, как и целые числа, образуют *кольцо* — алгебраическую структуру с операциями сложения, вычитания и умножения (без деления).

Для конкретизации структуры регулятора воспользуемся условием реализуемости регулятора:

$$m_R \leq n_R; \quad (4.6)$$

для упрощения задачи примем степени равными: $m_R = n_R$. Тогда число неизвестных параметров регулятора равно $2n_R + 1$.

Из условия равенства коэффициентов полиномов A_3 и A^* имеем систему уравнений для определения коэффициентов полиномов A_R и B_R .

Степень полинома A_3 равна сумме степеней полиномов A_0 и A_R , т.е. порядок системы равен сумме порядков объекта и регулятора. Такой же должна быть и степень желаемого характеристического полинома A^* . В силу того что полиномы A_0 и A_R имеют единичные старшие коэффициенты, а степень полинома B_0B_R меньше степени полинома A_0A_R , старший коэффициент полинома A_3 также равен единице. Как видно из (4.4), полином A^* имеет единичный старший коэффициент. Таким образом, из тождества (4.5) следует $n_0 + n_R$ уравнений.

Число неизвестных параметров должно равняться числу уравнений:

$$2n_R + 1 = n_0 + n_R;$$

откуда получим порядок регулятора:

$$n_R = n_0 - 1. \quad (4.7)$$

Порядок системы равен

$$n = n_0 + n_R = 2n_0 - 1;$$

таково же и число неизвестных параметров регулятора при выполнении условия (4.6).

Далее записывается система уравнений относительно искомым параметров. Матрица системы формируется из коэффициентов полиномов A_0 и B_0 ; она оказывается так называемой матрицей Сильвестра. Ее определитель — *результант* полиномов A_0 и B_0 — отличен от нуля, если полиномы взаимно просты. Таким образом, задача размещения корней разрешима, если характеристика вход-выход объекта является полной.

Пример 1. Рассмотрим задачу стабилизации маятника (см. рис. 1.7, а) в верхнем положении равновесия. Пусть масса сосредоточена в точке, трение и сопротивление среды отсутствуют. Малые отклонения маятника от верхнего положения равновесия описываются дифференциальным уравнением второго порядка

$$(p^2 - 1)y(t) = u(t),$$

где u — управляющее воздействие. Для простоты принято, что значение длины маятника численно равно значению ускорения силы тяжести.

Один из корней $p_{1,2} = \pm 1$ характеристического полинома дифференциального уравнения объекта положителен. Действительно, верхнее положение маятника не устойчиво.

Поскольку порядок объекта равняется двум, то из (4.7) найдем порядок (структуру) регулятора: $n_R = 1$. Следовательно, необходимо определить три коэффициента дифференциального уравнения регулятора

$$(p + a_{0R})u(t) = -(b_{1R}p + b_{0R})y(t).$$

Порядок системы равен трем; назовем три желаемых корня: $p_1^* = -1$; $p_2^* = -2$; $p_3^* = -4$ и сформируем в соответствии с выражением (4.4) желаемый характеристический полином:

$$A^*(p) = (p+1)(p+2)(p+4) = p^3 + 7p^2 + 14p + 8.$$

Один из корней характеристического полинома объекта (левый) оставляем на месте. Характеристический полином замкнутой системы в соответствии с уравнением (4.3) для данного примера имеет вид:

$$A_3(p) = (p^2 - 1)(p + a_{0R}) + 1 \cdot (b_{1R}p + b_{0R}) = p^3 + a_{0R}p^2 + (b_{1R} - 1)p + (b_{0R} - a_{0R}).$$

Из тождества $A_3(p)$ и $A^*(p)$ получим систему уравнений

$$a_{0R} = 7;$$

$$b_{1R} - 1 = 14;$$

$$b_{0R} - a_{0R} = 8,$$

решая которую получим искомые коэффициенты регулятора, обеспечивающего заданное расположение корней: $a_{0R} = 7$; $b_{1R} = 15$; $b_{0R} = 15$.

Передаточная функция разомкнутого контура стабилизации

$$W_p(s) = W_0(s)W_{R1}(s) = \frac{1}{s^2 - 1} \cdot \frac{15(s+1)}{s+7}$$

имеет диполь $z_1 = p_1 = -1$, равный оставляемому на месте корню объекта.

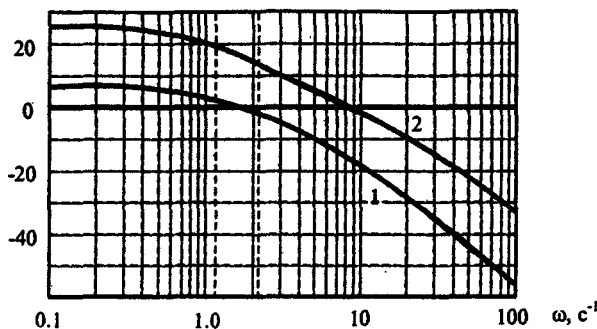


Рис. 4.5. ЛАЧХ к примеру решения задачи стабилизации операторным методом

На рис. 4.5 изображена ЛАЧХ контура (кривая 1). Частота среза системы $\omega_{\text{ср}1} = 2 \text{ с}^{-1}$.

Изменим положение желаемых корней системы: $p_1^* = -4$; $p_2^* = -8$; $p_3^* = -16$. Удаление корней от мнимой оси отражает стремление к повышению быстродействия системы. При этом получим следующую передаточную функцию регулятора:

$$W_{R2}(s) = \frac{225s + 540}{s + 28}.$$

На рис. 4.5 (кривая 2) изображена ЛАЧХ разомкнутого контура соответствующей системы. Наблюдаем расширение полосы частот; частота среза $\omega_{\text{ср}2} = 8 \text{ с}^{-1}$.

На рис. 4.6, а приведены графики процессов $y(t)$ в первой и второй системах при следующих начальных условиях: $y(0) = 1$; $y'(0) = y''(0) = 0$. Видно, что во второй системе процесс затухает быстрее — маятник быстрее возвращается в положение равновесия. Длительность процессов: $t_{p1} \approx 5 \text{ с}$; $t_{p2} \approx 1,25 \text{ с}$. На рис. 4.6, б изображены графики изменения управляющих воздействий $u(t)$; во второй системе уровни воздействий на объект гораздо выше: $|u_{\text{max}1}| \approx 20$; $|u_{\text{max}2}| \approx 200$.

Подбором полиномов B_R и A_R можно получить любой желаемый характеристический полином системы и даже добиться понижения степени за счет взаимного уничтожения старших коэффициентов. При этом часть корней полинома уходит в бесконечность. Поскольку неточная компенсация может дать полиномы с малыми отрицательными коэффициентами, часть корней переходит в правую полуплоскость. Системы, полученные таким образом, оказываются негрубыми — при малейшей неточности в реализации регулятора или несоответствии объекта модели система будет катастрофически неустойчивой — характеристический полином будет иметь большие по модулю правые корни.

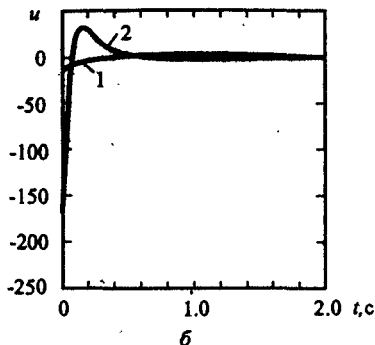
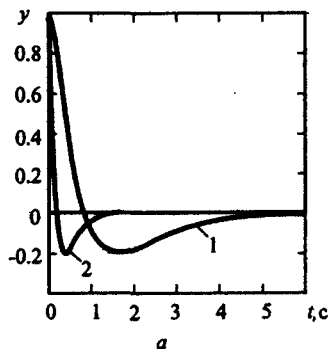


Рис. 4.6. Графики процессов к примеру решения задачи стабилизации операторным методом

4.3.2. Размещение собственных значений матрицы дифференциальных уравнений в форме пространства состояний

Пусть объект описан дифференциальными уравнениями в форме пространства состояний

$$\frac{dv}{dt} = Av + Bu. \quad (4.8)$$

Анализ выявил, что среди собственных значений матрицы A есть правые и/или сильно колебательные. Следовательно, такие собственные значения необходимо переместить.

В предположении, что измеряются все переменные состояния, скалярное управляющее воздействие формируется как их линейная функция:

$$u = -(k_1 v_1 + \dots + k_n v_n) = -Kv, \quad (4.9)$$

где K — матрица-строка. Соответствующее управляющее устройство имеет predetermined структуру; оно безынерционно, следовательно, не повышает порядка системы. Здесь решается задача параметрического синтеза — определения значений элементов матрицы обратной связи по состоянию K .

Дифференциальные уравнения системы получаются в результате подстановки выражения (4.9) в уравнение (4.8):

$$\frac{dv}{dt} = (A - BK)v. \quad (4.10)$$

Матрица замкнутой системы $A - BK$ должна иметь заданные собственные значения $\{p^*, i = 1, \dots, n\}$. Сформируем желаемую сопровождающую матрицу:

$$A^* = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0^* & -a_1^* & -a_2^* & \dots & -a_{n-1}^* \end{bmatrix},$$

в которой элементами последней строки являются коэффициенты желаемого характеристического полинома (4.4) с обратными знаками.

Пусть объект полностью управляем; примем, что уравнение (4.8) записано в управляемом каноническом базисе — матрица A имеет форму Фробениуса, а матрица-столбец B состоит из нулей кроме единицы в последней строке

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{bmatrix}.$$

Тогда матрица системы также имеет форму Фробениуса

$$A - BK = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ a_{n1} - k_1 & a_{n2} - k_2 & a_{n3} - k_3 & \dots & a_{nn} - k_n \end{bmatrix}.$$

Искомые коэффициенты регулятора легко находятся из равенства матриц A^* и $A - BK$:

$$k_i = a_{ni} + a_{i-1}^*; \quad i = 1, \dots, n. \quad (4.11)$$

В случае обратной связи по состоянию порядок системы совпадает с порядком объекта. Но это не говорит о простоте технического решения задачи стабилизации — измерение переменных состояния часто является проблемой.

Пример 2. Рассмотрим ту же задачу стабилизации верхнего положения равновесия маятника. Примем следующие переменные состояния: $v_1 = y$; $v_2 = y'$ и запишем дифференциальные уравнения маятника в управляемой канонической форме пространства состояний с матрицами

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (4.12)$$

Собственные значения матрицы объекта A равны ± 1 .

Назначим желаемые собственные значения матрицы системы: $p_1^* = -1; p_2^* = -2$. По характеристическому полиному

$$A^*(p) = (p+1)(p+2) = p^2 + 3p + 2$$

построим желаемую сопровождающую матрицу

$$A^* = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}.$$

Коэффициенты обратной связи по состоянию

$$u = -(k_1 v_1 + k_2 v_2)$$

найдем из условия равенства матриц $A - BK$ и A^*

$$A - BK = \begin{bmatrix} 0 & 1 \\ 1 - k_1 & -k_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}.$$

Значения параметров регулятора равны: $k_1 = 3; k_2 = 3$. Получен так называемый пропорционально-дифференцирующий (ПД) закон регулирования (см. п. 1.5.2):

$$u = -3(v_1 + v_2) = -\left(3y + 3\frac{dy}{dt}\right).$$

Передающая функция регулятора

$$W_{R1}(s) = 3(s+1)$$

не удовлетворяет условию реализуемости — степень полинома числителя превышает степень полинома знаменателя. Однако в данном случае реализуется регулятор состояния (4.9), который формирует управляющее воздействие по измеренным значениям производной $v_2 = dy/dt$, т.е. операция дифференцирования не используется.

Назначим теперь другие желаемые собственные значения: $p_1^* = -4; p_2^* = -8$. По соотношениям (4.11) получаются следующие значения коэффициентов обратных связей по состоянию: $k_1 = 33; k_2 = 12$.

На рис. 4.7 изображены ЛАЧХ для двух рассматриваемых примеров, построенные по передаточным функциям систем, разомкнутых относительно переменной управления:

$$W_{p1}(s) = \frac{3s+3}{s^2-1};$$

$$W_{p2}(s) = \frac{12s+33}{s^2-1}.$$

Видно, что требование большего быстродействия дает систему с более широкой полосой пропускания: $\omega_{cp1} \approx 3 \text{ с}^{-1}; \omega_{cp2} \approx 12 \text{ с}^{-1} = 4\omega_{cl}$.

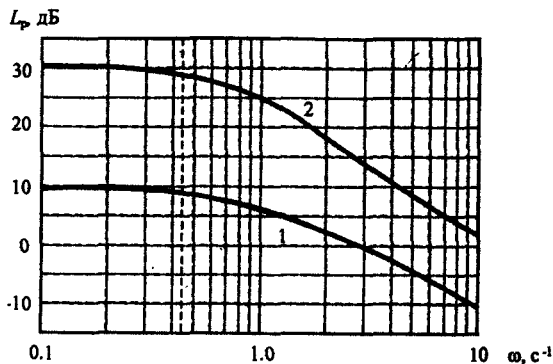


Рис. 4.7. ЛАЧХ к примеру решения задачи стабилизации методом пространства состояний

На рис. 4.8, а, б приведены кривые процессов в замкнутых системах с меньшим и большим быстродействием. Можно видеть, что во втором случае процесс стабилизации протекает в четыре раза быстрее: $t_{p1} \approx 5$ с; $t_{p2} \approx 1,25$ с $= t_{p1}/4$. Вместе с тем, стремление повысить быстродействие приводит к необходимости оказания на объект больших по уровню управляющих воздействий: $u_{1\max} \approx 3$; $u_{2\max} \approx 30$. Может оказаться, что такие управляющие воздействия нельзя реализовать из-за ограничений ресурсов. Кроме того, при больших отклонениях переменных линейные модели не адекватны реальной системе. Расширение полосы частот системы может потребовать учета малых постоянных времени, например, уточнения динамических свойств исполнительного механизма и измерительно-преобразовательных элементов.

Произвольный выбор желаемых корней и простота определения значений коэффициентов обратных связей по соотношениям (4.11) может привести к неверному выводу о том, что в замкнутой системе можно добиться любого качества процессов управления. В рамках линей-

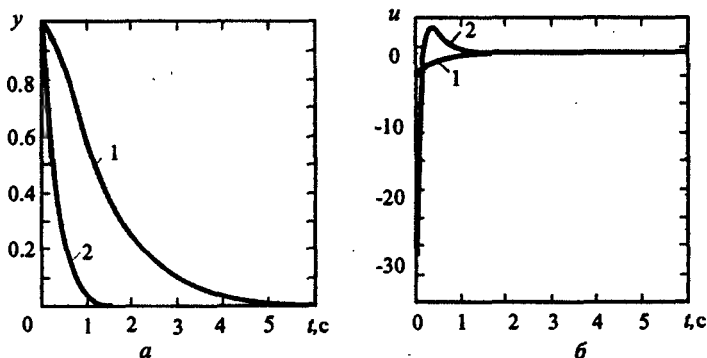


Рис. 4.8. К примеру решения задачи стабилизации методом пространства состояний. Графики процессов

ных математических моделей это, разумеется, так. Однако линейные модели адекватны реальным системам только для малых отклонений переменных состояния и управления и ограниченных диапазонов частот (см. рис. 2.21). Стремление к быстрому затуханию процессов — выбор больших по модулю желаемых корней, т.е. увеличение степени устойчивости (быстродействия) η , приводит к тому, что некоторые из переменных состояния и переменная управления за время процесса изменяются с большой скоростью и принимают очень большие значения. Для объяснения быстрых движений исходные модели оказываются не вполне адекватными системе — в них не учтены малые инерционности объекта, измерителей, исполнительных механизмов, ставшие теперь существенными. Поэтому при назначении желаемых собственных значений матрицы системы следует ориентироваться на границы области адекватности ($|p^*| < \Omega$). Кроме того, необходим анализ процессов в синтезированной системе при типовых и других начальных условиях с целью проверки допустимости отклонений переменных состояния $v(t)$ и управления $u(t)$.

Проблема выбора желаемых корней — это и есть основная проблема синтеза в описанной методике. Она должна решаться с учетом комплекса условий. При значительной априорной неопределенности о поведении объекта и системы задача решается путем итераций, и, по существу, результатом синтеза одновременно являются коэффициенты обратных связей и окончательно установленные желаемые корни.

4.3.3. Аналитическое конструирование регуляторов

Требования устойчивости и качества процессов можно описывать в неявной форме как экстремали тех или иных функционалов. Наиболее часто, в силу относительной простоты вычислений при достаточной физичности, применяют интегральные квадратичные функционалы (см. п. 3.11.2). В случае, когда объект описан в форме пространства состояний, интегральный квадратичный функционал записывается в виде

$$I = \int_0^{\infty} (v^T Q v + r u^2) dt, \quad (4.13)$$

где v — вектор состояния; u — скалярное управление; Q — неотрицательно-определенная весовая матрица; $r > 0$ — весовой коэффициент. Безусловная экстремаль $v^*(t)$ функционала (4.13) отвечает желаемому поведению и зависит от выбора весовых коэффициентов.

В п. 3.11.2 рассматривался вопрос о взаимосвязи интегральных и временных показателей качества. На примере улучшенной интегральной квадратичной оценки было показано, что большие значения коэффициента веса при производной процесса означают стремление к медленно затухающим экспоненциальным процессам с малыми значениями скоростей. Введение в функционал старших производных позволяет учесть большее разнообразие требований к протеканию процессов. Дополнение функционала членом $ru^2(t)$ означает косвенное ограничение энергии управления.

Задачей синтеза является определение матрицы коэффициентов обратной связи по состоянию $\mathbf{K}$ (см. п. 4.3.2), доставляющей минимум функционалу (4.13). Минимизация (4.13) при динамических ограничениях в виде дифференциальных уравнений объекта (4.8) дает условную экстремаль.

Оказывается (см. § 20.3), что матрица $\mathbf{K}$ коэффициентов обратных связей находится из соотношения

$$\mathbf{K} = \mathbf{B}^T \mathbf{P} / r,$$

где матрица $\mathbf{P}$ является решением нелинейного матричного уравнения Риккати (J. Riccati)

$$\mathbf{P}\mathbf{A} + \mathbf{A}^T \mathbf{P} - \mathbf{P}\mathbf{B}\mathbf{B}^T \mathbf{P} / r + \mathbf{Q} = 0. \quad (4.14)$$

Уравнения такого вида решаются численно.

Пример 3. Рассмотрим задачу стабилизации перевернутого маятника, заданного уравнениями состояний (4.8) с матрицами (4.12). Примем весовую матрицу $\mathbf{Q}$ в функционале (4.13) единичной; положим также $r = 1$. В результате проведения необходимых вычислений, включающих решение уравнения Риккати (4.14), находим матрицу обратных связей:

$$\mathbf{K} = (2,41 \quad 2,41).$$

При этом собственные значения замкнутой матрицы системы получаются следующими: $p_1 = -1$; $p_2 = -1,41$.

Поскольку при значительной исходной неопределенности трудно сразу выбрать весовые коэффициенты функционала, задача решается путем последовательных приближений.

Изменим весовой коэффициент r функционала (4.13) при управлении, приняв $r = 10$. Соответствующая матрица обратных связей будет равна:

$$\mathbf{K} = (2,05 \quad 2,05),$$

а собственные значения матрицы замкнутой системы получатся следующими: $p_1 = -1$; $p_2 = -1,05$.

Метод аналитического конструирования регуляторов для многомерных объектов будет подробно изложен далее.

4.3.4. Синтез наблюдателя состояний

При построении регуляторов в п. 4.3.3 предполагалось, что все переменные состояния объекта управления могут быть измерены непосредственно. Однако, как правило, измеряются только переменные выхода, число которых меньше порядка модели объекта. Уравнения состояния (4.8) в этом случае дополняются уравнением выхода:

$$\begin{aligned}\frac{dv}{dt} &= Av + Bu; \\ y &= Cv.\end{aligned}\tag{4.15}$$

Если объект наблюдаем полностью, то по измеренным значениям переменной выхода y можно вычислять текущее состояние объекта. При этом управляющее воздействие на объект (4.15) формируется по оценкам вектора состояния

$$u = -K\hat{v}.$$

Наблюдатель состояния представляет собой модель объекта, охваченную обратной связью по отклонению δy выходов модели $\hat{y}$ и объекта y (рис. 4.9).

Оценка $\hat{v}$ вектора состояния отличается от состояния v объекта из-за различия начальных условий, действующих на объект возмущений, а также неточности описания объекта. Однако при правильном выборе матрицы обратной связи наблюдателя L оценка $\hat{v}$ должна асимптотически стремиться к состоянию объекта.

Можно показать, что задача синтеза наблюдателя — определения матрицы L — является дуальной по отношению к задаче синтеза регулятора — определения матрицы K . Поэтому матрицу наблюдателя L можно найти теми же методами, если вместо пары матриц (A, B) принять пару (A^T, C^T) . При назначении желаемых собственных значений

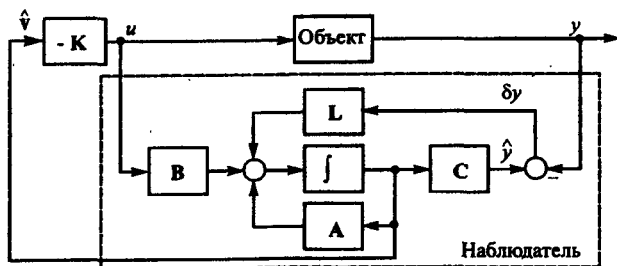


Рис. 4.9. Наблюдатель состояния

матрицы или весовых коэффициентов функционалов вида (4.13) необходимо стремиться к большему быстрдействию контура наблюдателя.

Пример 4. Рассмотрим задачу синтеза наблюдателя для перевернутого маятника. Пусть измеряется положение маятника y . Тогда, в силу выбора переменных состояния $y = v_1$; $y' = v_2$, матрица выхода $C = (1 \ 0)$.

Проведем синтез методом размещения корней. Назначим собственные значения наблюдателя значительно дальше от мнимой оси, чем собственные значения системы, т. е. примем $p_1^* = -10$; $p_2^* = -20$. Получим следующую матрицу наблюдателя $L = (30 \ 201)^T$.

Наблюдатель вместе с регулятором состояния образуют динамический регулятор, входом которого является выход объекта, а выходом — управляющее воздействие на объект (см. рис. 4.9). Передаточная функция этого регулятора для рассматриваемого примера равна:

$$W_R(s) = \frac{693(s+1)}{s^2 + 33s + 293}.$$

Оказалось, что нуль передаточной функции регулятора $z_1 = -1$ равен левому, оставляемому на месте, полюсу передаточной функции объекта. Корни характеристического полинома замкнутой системы

$$A_3(s) = s^4 + 33s^3 + 292s^2 + 660s + 400$$

в точности равны назначенным корням системы: $p_1 = -1$; $p_2 = -2$ и наблюдателя $p_3 = -10$; $p_4 = -20$.

Отметим в заключение, что недостатком синтезированного наблюдателя является его избыточность. Поскольку одна из переменных состояния $v_1 = y$ измеряется, следует синтезировать наблюдатель, порядок которого меньше порядка модели объекта (см. § 20.2).

4.4. СИНТЕЗ СИСТЕМ УПРАВЛЕНИЯ, ИНВАРИАНТНЫХ К ВОЗМУЩЕНИЯМ

Инвариантность управляемой переменной к возмущениям — одно из основных свойств систем управления.

Если о возмущении имеется полная текущая, а о модели объекта — полная априорная информация, то теоретически возможна его полная компенсация — достижение абсолютной инвариантности переменной выхода объекта к возмущению. Условия абсолютной инвариантности для типовых структур приведены в § 3.7, и здесь их можно не повторять. Реализация условий абсолютной инвариантности к непосредственно измеряемым воздействиям возможна, однако часто наталкивается на проблему физической осуществимости передаточной функции канала компенсации.

Если о возмущениях нет полной информации, то создаются системы с контурами обратных связей. В системах, реализующих принцип отрицательной обратной связи, управляющее воздействие формируется по информации о следствии возмущений — по отклонениям переменной выхода объекта.

Практически стремятся обеспечить селективную инвариантность — независимость или слабую зависимость установившихся процессов от воздействий с заданным спектром. Условием селективной абсолютной инвариантности является равенство нулю передаточной функции системы на полюсах $\{s_f\}$ изображения возмущения $F(s)$, т. е.

$$\Phi_{yf}(s_f) = 0.$$

Селективная инвариантность до ϵ имеет место, если

$$|\Phi_{yf}(s_f)| < \epsilon \neq 0.$$

Обычно имеется априорная информация о частотном спектре возмущения, например, известно, что его частоты не выше Ω_f . Для селективной инвариантности до ϵ к таким возмущениям должно быть обеспечено условие

$$\forall \omega \leq \Omega_f : |\Phi_{yf}(j\omega)| < \epsilon.$$

Рассмотрим вопросы выбора структур операторов управляющих устройств и значений их параметров, т. е. синтеза алгоритмов управляющих устройств селективно инвариантных замкнутых систем.

4.4.1. Синтез статических систем

На рис. 4.10 изображена структурная схема системы управления с обратной связью. Здесь обсуждается задача синтеза по требованиям к установившимся постоянным значениям переменных. Поэтому рассматривается частная модель для равновесных режимов, когда все звенья структурной схемы статические и заданы их коэффициенты усиления. Управляющее воздействие u^* рассчитано из условия равенства выхода объекта заданному постоянному значению y^* :

$$u^* = \frac{1}{k_u} y^*.$$

Если к объекту приложено постоянное по уровню возмущение f_o , то появляется отклонение управляемой переменной от заданного значения:

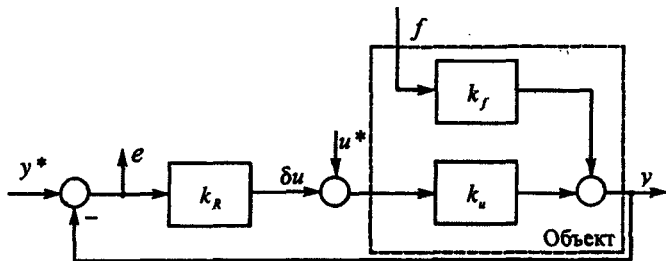


Рис. 4.10. Статическая система

$$y = y^* + k_f f_0.$$

В том случае, когда существенное возмущение нельзя скомпенсировать, его влияние на управляемую переменную можно ослабить с помощью отрицательной обратной связи, т. е. путем создания замкнутой системы управления. Это предполагает измерение фактического значения управляемой переменной, его сравнение с заданием и выявление отклонения (ошибки):

$$e = y^* - y = -k_f f_0, \quad (4.16)$$

возникающего как следствие возмущения. Дополнительное управляющее (регулирующее) воздействие на объект формируется в зависимости от величины этого отклонения

$$\delta u = k_R e.$$

Такой закон регулирования называют *пропорциональным* (см. п. 1.5.2). Значение ошибки в системе с обратной связью (см. рис. 4.10) равно:

$$e = \frac{-k_f}{1 + k_u k_R} f_0. \quad (4.17)$$

Сравнивая выражения (4.16) и (4.17) видим, что в системе с обратной связью отклонение управляемой переменной от задания уменьшается в $(1 + k_u k_R)$ раз. Влияние возмущения на управляемую переменную тем слабее, чем больше коэффициент усиления разомкнутого контура

$$k_p = k_u k_R.$$

Синтез статической системы по требованию к установившейся точности сводится к определению коэффициента усиления регулятора k_R . Пусть даны максимальное значение возмущения $f_{\max}$ и допустимая статическая ошибка $e_{\max}$. Из соотношения (4.17) имеем:

$$\frac{k_f}{1 + k_u k_R} \dot{f}_{\max} \leq e_{\max}$$

следовательно, искомое значение коэффициента усиления регулятора выбирается с учетом условия

$$k_R \geq \frac{1}{k_u} \left(\frac{k_f f_{\max}}{e_{\max}} - 1 \right).$$

4.4.2. Синтез астатических систем по требованию к точности подавления степенных возмущений

В системе с обратной связью (см. рис. 4.10) выберем интегральный закон регулирования:

$$\delta u = k_I \int_0^t e(\tau) d\tau.$$

Передаточная функция такого регулятора запишется так:

$$W_R(s) = \frac{k_I}{s}.$$

Изображение по Лапласу постоянного возмущения среды на объект имеет вид

$$F(s) = \frac{f_0}{s}.$$

Таким образом, передаточная функция регулятора имеет полюс, равный полюсу изображения возмущения, т. е. в контуре присутствует «модель среды», что обеспечивает селективную абсолютную инвариантность (см. п. 3.7.3) — нулевую установившуюся ошибку.

Звено интегрирующего типа является *астатическим* — при постоянном ненулевом значении входа переменная выхода не остается постоянной — она растет или убывает с постоянной скоростью до бесконечности. Можно сказать, что такое звено имеет бесконечное усиление по постоянному сигналу, что и обеспечивает нулевую установившуюся ошибку замкнутой системы.

Рассмотрим случай, когда возмущение изменяется с постоянной скоростью

$$f(t) = f_1 t.$$

$$F(s) = \frac{f_1}{s^2}$$

имеет двукратный нулевой полюс. Поскольку это на единицу больше кратности нулевого полюса передаточной функции регулятора, установившаяся ошибка отлична от нуля и равна

$$\begin{aligned} e_{\text{уст}} &= \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} s\Phi_{ef}(s)F(s) = \\ &= \lim_{s \rightarrow 0} s \frac{sk_f}{s + k_u k_l} \frac{f_1}{s^2} = \frac{k_f}{k_u k_l} f_1. \end{aligned}$$

Здесь имеет место селективная инвариантность до ϵ . Легко найти значение параметра настройки регулятора, если заданы максимальное значение скорости изменения возмущения f_{imax} и допустимая установившаяся ошибка e_{max}

$$k_l \geq \frac{k_f f_{\text{imax}}}{k_u e_{\text{max}}}.$$

Рассмотрим теперь общий случай степенных воздействий (2.49) изображениями, имеющими $(l + 1)$ -кратный нулевой полюс.

В соответствии с принципом внутренней модели выбирается передаточная функция регулятора — внутреннего компенсатора возмущения:

$$W_R(s) = \frac{k_v}{s^v},$$

также имеющего нулевые полюсы.

Задачей структурного синтеза алгоритма управления является определение кратности нулевого полюса передаточной функции регулятора и задача параметрического синтеза — вычисление коэффициента передаточной функции k_v . Положим, что передаточные функции объекта по обоим каналам $W_u(s)$ и $W_f(s)$ не имеют нулевых полюсов и нулей, т. е. $W_u(0) = k_u$; $W_f(0) = k_f$.

Как следует из п. 3.7.3, установившаяся ошибка зависит от соотношения степени воздействия l и порядка астатизма v :

- при $v \geq l + 1$ ошибка равна нулю (имеет место селективная абсолютная инвариантность при любых ненулевых значениях параметров k_u, k_f, k_v);
- при $v < l$ ошибка стремится к бесконечности (система не работоспособна);

• при $v = l \neq 0$ ошибка постоянна и зависит от параметров объекта и регулятора:

$$e_{\text{уст}} = \frac{k_f}{k_u k_v} f_l.$$

Отсюда выбирается структура внутреннего компенсатора возмущения. Если требуется нулевая установившаяся ошибка, то $v = l + 1$; при этом параметр k_v необходимо конкретизировать исходя из дополнительных условий. Если же установившаяся ошибка не должна превышать заданного значения e_{max} , то $v = l$, а параметр настройки выбирается из условия

$$k_v \geq \frac{k_f f_{l\text{max}}}{k_u e_{\text{max}}}.$$

4.4.3. Синтез систем по требованию к точности подавления гармонических возмущений

Если известно, что к объекту приложено гармоническое возмущающее воздействие $f(t)$ с частотой ω_f и амплитудой a_f , то для селективной абсолютной инвариантности передаточная функция управляющего устройства должна иметь полюсы $\pm j\omega_f$, т.е. содержать консервативное звено (см. п. 3.6.3). Тогда при любой амплитуде возмущения a_f амплитуда установившейся реакции системы a_y будет равна нулю.

В том случае, когда контур управления не содержит консервативного звена на частоте возмущения, имеет место селективная инвариантность до ϵ . Амплитуда установившейся реакции системы a_y не равна нулю и зависит от амплитуды возмущения a_f и параметров системы:

$$a_y = |\Phi_{\text{ш}}(j\omega_f)| a_f = \left| \frac{W_{\text{of}}(j\omega_f)}{1 + W_{\text{om}}(j\omega_f)W_R(j\omega_f)} \right| a_f. \quad (4.18)$$

Чем больше усиление контура на частоте возмущения

$$|W_p(j\omega_f)| = |W_{\text{om}}(j\omega_f)W_R(j\omega_f)| \gg 1,$$

тем меньше амплитуда установившейся реакции.

Положим, что задана допустимая амплитуда реакции на выходе объекта a_{ymax} . Тогда необходимое усиление управляющего устройства находится так:

$$\left| W_R(j\omega_f) \right| \geq \frac{\left| W_{of}(j\omega_f) \right| a_f}{\left| W_{ou}(j\omega_f) \right| a_{y\max}}.$$

Если имеется априорная информация о низкочастотном характере возмущений и известно, что частоты существенных сигналов не превосходят Ω_f , то усиление контура должно быть не ниже требуемого значения на всем интервале частот $[0, \Omega_f]$. Выполнение этого условия легко контролируется по логарифмическим частотным характеристикам.

Пусть амплитуда возмущений должна быть ослаблена не менее чем в сто раз для всех частот из интервала $[0, \Omega_f]$:

$$a_y / a_f \leq 0,01.$$

Тогда из (4.18) получим

$$\forall \omega \leq \Omega_f : L_{of}(\omega) - L_{ou}(\omega) - L_R(\omega) \leq -40 \text{ дБ},$$

или

$$L_R(\omega) > L_{of}(\omega) - L_{ou}(\omega) + 40 \text{ дБ}.$$

Если возмущение приложено к выходу объекта, то

$$L_{of}(\omega) \equiv 0,$$

а требуемое усиление разомкнутого контура

$$\forall \omega \leq \Omega_f : L_p(\omega) = L_R(\omega) + L_{ou}(\omega) > 40 \text{ дБ}.$$

Таким образом, требования установившейся точности определяют низкочастотную часть ЛАЧХ разомкнутой системы, которые имеют низкочастотную асимптоту с наклоном -20ν дБ/дек.

4.5. СИНТЕЗ СЛЕДЯЩИХ СИСТЕМ ИЗ УСЛОВИЯ ТРЕБУЕМОЙ ТОЧНОСТИ ВОСПРОИЗВЕДЕНИЯ

Говоря о точности следящих систем, прежде всего, имеют в виду малую установившуюся ошибку воспроизведения задающего воздействия. Примем за выход следящей системы с единичной отрицательной обратной связью (рис. 4.11) переменную ошибки $e(t)$. Тогда задачей синтеза следящей системы, в первую очередь, является определение передаточной функции $W_g(s)$ внутреннего компенсатора воздействия, обеспечивающего селективную инвариантность переменной $e(t)$ к воздействию $g(t)$. При решении этой задачи сохраняются в силе основные положения синтеза систем, инвариантных к возмущению (см. § 4.4).

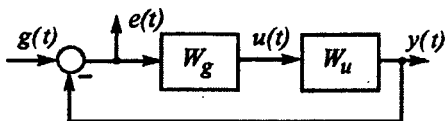


Рис. 4.11. Структура следящей системы

Пусть задающие воздействия моделируются степенными функциями времени. Правила выбора порядка астатизма приведены в § 4.4. При реализации внутреннего компенсатора необходимо учесть, что объекты управления следящих систем обычно астатические (например, двигатели с редукторами). Поэтому при выборе структуры внутреннего компенсатора, в частности, числа нулевых полюсов передаточной функции $W_g(s)$ необходимо учесть наличие таких полюсов у передаточной функции $W_u(s)$ объекта по каналу управления.

Пусть задающие воздействия имеют вид гармонических функций времени с амплитудами a_g и частотами ω_g , не превосходящими заданных значений:

$$a_g \leq A_g; \quad \omega_g \leq \Omega_g.$$

Амплитуды установившихся ошибок равны:

$$a_e = |\Phi_{eg}(j\omega_g)| a_g,$$

где

$$\Phi_{eg}(j\omega) = \frac{1}{1 + W_u(j\omega)W_g(j\omega)} = \frac{1}{1 + W_p(j\omega)}$$

представляет собой частотную характеристику замкнутой системы по ошибке. Условие селективной инвариантности до ε — требуемая точность воспроизведения в установившихся режимах — выражается так:

$$a_e \leq A_e,$$

где A_e — допустимое значение амплитуды ошибки. Требование удовлетворяется при условии:

$$\forall \omega \leq \Omega_g : |\Phi_{eg}(j\omega)| \leq A_e / A_g.$$

Поскольку на частотах воспроизводимых сигналов усиление разомкнутого контура много больше единицы:

$$|W_p(j\omega)| \gg 1,$$

то последнее условие запишется так:

$$\forall \omega \leq \Omega_g : |W_p(j\omega)| \leq A_g / A_e.$$

То же условие для ЛАЧХ имеет вид

$$\forall \omega \leq \Omega_g : L_p(\omega) > 20 \lg(A_g / A_e).$$

Графически это условие интерпретируется как требование прохождения ЛАЧХ выше запретной области на низких частотах (рис. 4.12).

На рис. 4.13 изображен пример низкочастотной части ЛАЧХ разомкнутого контура астатической системы. Если задающее воздействие имеет вид гармонического сигнала с фиксированной частотой ω_g , то для обеспечения требуемой точности воспроизведения ЛАЧХ контура должна проходить выше контрольной точки с координатами $(\omega_g, L_x = 20 \lg(A_g / A_e))$, как это показано на рис. 4.13.

Положим теперь, что следящая система должна воспроизводить задающие воздействия, о которых известно лишь, что скорости их изменения не превосходят $g_{1\max}$, а ускорения — $g_{2\max}$. В этом случае целесообразно рассматривать некоторый эквивалентный гармонический сигнал [13]

$$g_s(t) = a_s \sin \omega_s t,$$

амплитуда a_s и частота ω_s , которого подбираются из условий

$$\max_t \dot{g}_s(t) = a_s \omega_s = g_{1\max};$$

$$\max_t \ddot{g}_s(t) = a_s \omega_s^2 = g_{2\max}.$$

Отсюда определяются параметры эквивалентного гармонического сигнала

$$\begin{aligned} \omega_s &= \frac{g_{2\max}}{g_{1\max}}, \\ a_s &= \frac{g_{1\max}^2}{g_{2\max}}. \end{aligned} \quad (4.19)$$

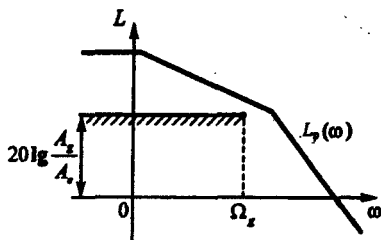


Рис. 4.12. Графическая иллюстрация условия селективной инвариантности к гармоническим воздействиям

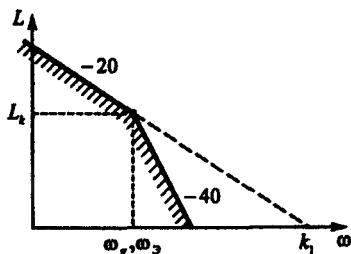


Рис. 4.13. ЛАЧХ разомкнутого контура астатической системы

По этим значениям можно построить контрольную точку (см. рис. 4.13) с координатами $(\omega_g, L_g = 20 \lg(a_g/A_g))$, выше которой должна проходить логарифмическая амплитудно-частотная характеристика разомкнутого контура для достижения требуемой точности воспроизведения.

Если рассматривать задающие воздействия, скорость которых равна максимальному значению $g_{1\max}$, а ускорение меньше $g_{2\max}$, то в соответствии с (4.19) частота ω , эквивалентного гармонического воздействия будет уменьшаться, а амплитуда — возрастать. При этом контрольная точка будет перемещаться влево по прямой с наклоном -20 дБ/дек. В предельном случае, когда $g_{2\max} = 0$, вместо гармонического сигнала получим линейное воздействие

$$g(t) = g_{1\max} t.$$

Для воспроизведения такого воздействия система должна обладать астатизмом первого порядка, т.е. наклон низкочастотной асимптоты ЛАЧХ должен быть равен -20 дБ/дек. Для достижения требуемой точности слежения добротность контура по скорости k_1 должна быть

$$k_1 \geq g_{1\max} / A_g.$$

Если амплитуда установившейся ошибки воспроизведения сигнала фиксированной частоты ω_g должна равняться нулю, т.е. требуется селективная абсолютная инвариантность, то передаточная функция компенсатора $W_g(S)$ должна иметь мнимые полюсы $\pm j\omega_g$, т.е. компенсатор должен содержать консервативное звено.

В заключение отметим еще раз, что условия инвариантности ошибки следящей системы к сигналу задания те же, что и условия инвариантности выхода системы к возмущению, приведенному к выходу объекта. Астатизм и повышение усиления контура на низких частотах обеспечивают не только инвариантность до ϵ переменной ошибки к заданию (ковариантность выхода системы с заданием), но также и инвариантность к низкочастотным возмущениям. Более того, в силу высокого усиления на этих частотах условия инвариантности оказываются малочувствительными и к вариациям элементов системы.

4.6. КОРРЕКЦИЯ СИСТЕМ УПРАВЛЕНИЯ

Синтез систем управления из условий селективной инвариантности обеспечивает требуемую точность подавления возмущений и воспроизведения задания в установившихся режимах. На этом этапе синтеза не учитываются требования к устойчивости и свободным движениям при нену-

левых предначальных условиях, а также к переходным процессам, вызванным приложенными воздействиями (посленачальными условиями).

Если селективная инвариантность достигнута в системе без контуров — за счет образования новых путей передачи воздействий или преобразования старых, то задачи инвариантности и устойчивости решаются независимо. Все звенья в отдельности должны быть устойчивыми. В противном случае необходимо принять меры по их стабилизации.

Во многих практически важных случаях требуемая установившаяся точность отработки воздействий достигается образованием контуров с большим усилением на частотах воздействий. Как известно, наличие контуров приводит к тому, что собственные свойства системы могут существенно отличаться от свойств звеньев. В частности, система, контур которой образован устойчивыми звеньями, может оказаться неустойчивой и наоборот. В общем случае, чем выше усиление контура, образованного устойчивыми звеньями, тем больше проблем с устойчивостью замкнутой системы — условия инвариантности и устойчивости оказываются противоречивыми.

Как было показано в § 3.4, при замыкании контура с передаточной функцией W_p полюсы, образующие диполи, остаются на месте, т. е. в точности являются корнями характеристического полинома замкнутой системы. Если контур имеет малое усиление на собственных частотах отдельных звеньев, то характеристический полином имеет корни, приближенно равные полюсам передаточных функций этих звеньев. Следовательно, и в системах с контурами можно добиться определенной независимости (слабой зависимости) условий инвариантности и устойчивости. Рассмотрим пример такой системы.

Пусть в системе с передаточной функцией объекта (неизменяемой части) вида

$$W_0(s) = \frac{k_0}{(T_1s + 1)(T_2s + 1)(T_3s + 1)}$$

выполнено условие селективной абсолютной инвариантности к постоянному воздействию — передаточная функция управляющего устройства является интегратором

$$W_R(s) = \frac{k_{R1}}{s}.$$

В этой системе с астатизмом первого порядка установившаяся ошибка при постоянном воздействии равна нулю при любой отличной от нуля добротности контура по скорости

$$k_1 = k_0 k_{R1}.$$

Выберем значение k_1 таким малым, чтобы для всех полюсов $\{s_i = -1/T_i\}$ передаточной функции W_0 выполнялось условие малости усиления контура

$$|W_p(j|s_i|)| \ll 1,$$

т. е. практически имело место

$$L_p(|s_i|) = 20 \lg |W_p(j|s_i|)| < -20 \text{ дБ},$$

где $W_p(s)$ — передаточная функция разомкнутого контура. При замыкании такого контура полюсы передаточной функции $W_0(s)$ перемещаются мало — характеристический полином замкнутой системы имеет корни, приближенно равные $s_i = -1/T_i$. На рис. 4.14 изображена соответствующая асимптотическая ЛАЧХ $L_p(\omega)$.

Большое усиление на нулевой частоте постоянного воздействия и малое усиление на частотах устойчивых звеньев позволяют в этом частном случае системы с обратной связью достичь компромисса между инвариантностью и устойчивостью. Заметим, что на нулевой частоте — собственной частоте интегрирующего управляющего устройства — контур имеет бесконечно большое усиление. Поэтому при замыкании системы нулевой полюс перемещается влево вдоль действительной отрицательной полуоси до значения, приближенно равного $-k_1$. Переходные процессы в системе и свободные движения в основном определяются этим корнем. Система имеет малое быстроедействие.

В общем случае условия инвариантности приходят в противоречие с условиями устойчивости. Пусть к рассмотренной системе приложены не только постоянные, но и низкочастотные воздействия. Их воспроизведение или/и ослабление с требуемой точностью, т.е. селективная инвариантность до ϵ , требует повышения усиления контура на частотах $\omega < \Omega_f$. В данной системе этого можно добиться повышением добротности k_1 . Однако, как видно из рис. 4.14, при этом повышается усиление контура на собственных частотах объекта $|s_1|$, $|s_2|$ и $|s_3|$. При замыкании контура они значительно перемещаются, обуславливая чрезмерную колебательность процессов или даже неустойчивость системы.

Таким образом, в общем случае возникает задача стабилизации системы без нарушения условий инвариантности.

В принципе, для стабилизации неустойчивой системы можно применить известные методы стабилизации объектов. Однако они не учитывают ограничений, связанных с сохранением условий инвариантности. Поэтому традиционно ставится и решается задача *коррекции* — специального формирования характеристик разомкнутых контуров с

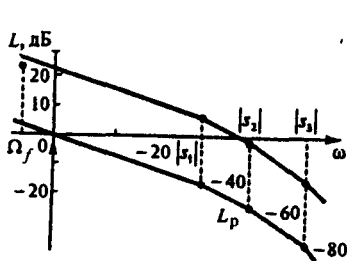


Рис. 4.14. Пример
асимптотической ЛАЧХ

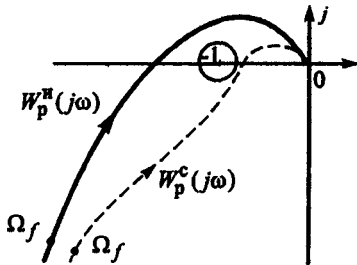


Рис. 4.15. АФХ контура до и после
коррекции

целью удовлетворения требований к устойчивости замкнутых систем и к переходным процессам без нарушения достигнутой ранее точности систем в установившихся режимах.

На рис. 4.15 непрерывной линией изображена АФХ $W_p^n(j\omega)$ контура инвариантной системы. Из критерия Найквиста следует, что при замыкании системы пара корней характеристического полинома перейдет через мнимую ось в правую полуплоскость — кривая охватывает точку $(-1, j0)$ по часовой стрелке. На том же рисунке пунктирной линией изображена характеристика $W_p^c(j\omega)$ скорректированной системы.

Если передаточная функция разомкнутой системы W_p минимально-фазовая (не имеет правых нулей и полюсов), то можно ограничиться построением только соответствующей амплитудно-частотной характеристики. Особенно удобны ЛАЧХ. На рис. 4.16 непрерывной и пунктирной линиями изображены ЛАЧХ L_p^n и L_p^c инвариантной системы до и после коррекции.

В типовых одноконтурных системах с единичной отрицательной обратной связью (см. рис. 4.3) передаточная функция разомкнутого контура

$$W_p(s) = B_p(s)/A_p(s)$$

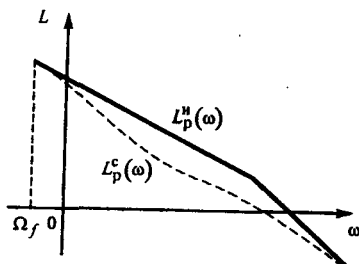


Рис. 4.16. ЛАЧХ контура до и после коррекции

однозначно определяет характеристический полином замкнутой системы

$$A_3(s) = A_p(s) + B_p(s),$$

а также передаточные функции по выходу

$$\Phi_y(s) = \frac{B_p(s)}{A_3(s)}$$

и по ошибке

$$\Phi_e(s) = \frac{A_p(s)}{A_3(s)}.$$

Поэтому, формируя желаемым образом передаточную функцию W_p , можно удовлетворить многообразие требований к замкнутой системе — воспроизведения и подавления низкочастотных воздействий, устойчивости, подавления высокочастотных помех, ослабления влияния вариаций операторов звеньев.

Формирование желаемой передаточной функции W_p^* , по существу, означает выбор рационального соотношения нулей и полюсов при заданном коэффициенте передачи контура. Для минимально-фазовых систем наиболее удобно такой выбор проводить на базе асимптотических ЛАЧХ, изломы которых соответствуют модулям нулей и полюсов. По ЛАЧХ непосредственно оцениваются усиления контура на частотах модулей и полюсов, а следовательно, можно приближенно судить о степени перемещения полюсов при замыкании контура с данным коэффициентом передачи.

Метод коррекции инвариантных систем в комплексно-частотной области сводится к обеспечению тождества передаточной функции скорректированного контура $W_p^c(s)$ желаемой

$$W_p^c(W^k) \equiv W_p^*, \quad (4.20)$$

где W^k — передаточная функция звена коррекции.

Основными этапами решения задачи коррекции являются:

- формирование желаемой передаточной функции разомкнутого контура;
- выбор места включения звена коррекции (топологический синтез);
- определение передаточной функции звена коррекции (структуры оператора и значений параметров), удовлетворяющей тождеству (4.20).

4.7. МЕТОДИКИ ФОРМИРОВАНИЯ ЖЕЛАЕМЫХ ПЕРЕДАТОЧНЫХ ФУНКЦИЙ

Синтез по требованиям селективной инвариантности дает систему, передаточная функция разомкнутого контура которой $W_p^*(s)$ имеет большие по модулю значения на спектре воздействий.

Желаемая передаточная функция $W_p^*(s)$ должна прежде всего, иметь те же полюсы воздействий, что и $W_p^n(s)$; это сохранит свойство селективной абсолютной инвариантности к этим воздействиям (в случае нулевых полюсов — астатизм). На частотах существенных воздействий модули желаемой передаточной функции должны приближенно равняться модулям исходной:

$$\forall \omega \in \Omega_f : |W_p^*(s)| \approx |W_p^n(s)|;$$

это сохранит свойство селективной инвариантности до ϵ . Поэтому коррекция контура означает формирование его передаточной функции вне спектра отрабатываемых воздействий, а именно, на более высоких частотах. Как указывалось ранее, повышение усиления контура на частотах $\omega \in \Omega_f$ приводит при замыкании контура к сильному перемещению тех полюсов передаточной функции $W_p(s)$, модули которых лежат в области Ω_f . Исключение составляют диполи. Если в области Ω_f нет нулей передаточной функции $W_p(s)$, то корни характеристического полинома замкнутой системы $A_3(s)$ будут большими по модулю, чем Ω_f . Таким образом, оказывается, что требования точности отработки непосредственно неизмеряемых воздействий определяют более низкочастотные характеристики контура по сравнению с собственными частотами системы.

В области высоких частот контур должен иметь малое усиление. Это способствует подавлению высокочастотных помех и обеспечивает невмешательство в область, где модель не адекватна описываемым элементам и объекту (область немоделируемой динамики). Малое усиление контура в области высоких частот приводит к тому, что характеристический полином замкнутой системы будет иметь корни, близкие к большим по модулю полюсам передаточной функции $W_p(s)$. Желаемая передаточная функция обычно содержит большие по модулю устойчивые полюсы исходной передаточной функции.

Таким образом, формирование желаемой передаточной функции сводится к выбору *средних* по модулю полюсов и нулей. Именно они определяют желаемые корни характеристического полинома и, в основном, вид переходных процессов.

В области средних частот усиления контура близки к единице, а ЛАЧХ пересекает ось абсцисс.

Выделим два основных подхода к формированию желаемой передаточной функции в среднечастотной области. Как известно, переходные процессы возникают из-за посленазначенных условий, вызванных воздействиями, приложенными в начальный момент времени (собственные сопровождающие составляющие вынужденных движений). Формирование только корней характеристического полинома означает синтез по требованиям к свободным движениям, а формирование соотношения левых корней характеристического полинома и нулей передаточной функции замкнутой системы — синтез по требованиям к переходным процессам по одной из пар вход-выход. Как было показано, приближение нулей передаточной функции к полюсам (образование диполей) уменьшает соответствующие составляющие переходного процесса при любых воздействиях. При синтезе следящих систем обеспечивают надлежащее взаимное расположение левых нулей и полюсов по каналу воспроизведения задания.

Типовому расположению корней характеристического полинома и нулей передаточной функции замкнутой системы отвечает типовое соотношение усиления, нулей и полюсов передаточной функции $W_p^*(s)$, а следовательно, типовой вид среднечастотного участка ЛАЧХ $L_p^*(\omega)$. Наиболее часто рассматриваются следующие типы передаточных функций:

$$W_p^*(s) = \frac{k_2(T_2s+1)}{s^2(T_3s+1)}; \quad (4.21)$$

$$W_p^*(s) = \frac{k_2(T_2s+1)}{s^2(T_3s+1)^2}; \quad (4.22)$$

$$W_p^*(s) = \frac{k_3(T_2s+1)^2}{s^3(T_3s+1)}; \quad (4.23)$$

$$W_p^*(s) = \frac{k_3(T_2s+1)^2}{s^3(T_3s+1)^2} \quad (4.24)$$

При $T_3 < T_2$. Им отвечают типовые асимптотические ЛАЧХ (рис. 4.17).

В соответствии с выражениями (4.21) — (4.24) все ЛАЧХ образуются тремя отрезками прямых, имеющих наклоны:

−40, −20, −40 дБ/дек, (2−1−2);

−40, −20, −60 дБ/дек, (2−1−3);

−60, −20, −40 дБ/дек, (3−1−2);

−60, −20, −60 дБ/дек, (3−1−3).

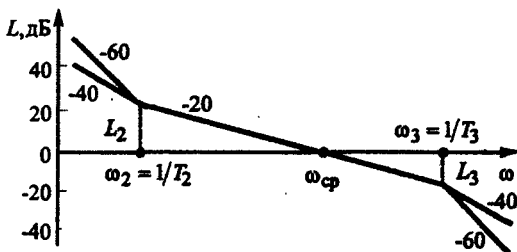


Рис. 4.17. Типовые асимптотические ЛАЧХ

Средняя асимптота имеет наклон -20 дБ/дек, а левая и правая — различные наклоны, определяемые числом сильно перемещаемых малых по модулю полюсов и слабо перемещаемых больших по модулю полюсов передаточной функции разомкнутой системы.

Типовые ЛАЧХ характеризуются частотой среза $\omega_{\text{ср}}$ и двумя параметрами L_2 и L_3 . Все они, безусловно, обеспечивают устойчивость замкнутой системы и достаточно хорошее относительное затухание переходных процессов. Частота среза определяет масштаб на плоскости корней — среднеегеометрическое корней, имеющих типовое расположение. Чем больше $\omega_{\text{ср}}$, тем больше модули корней и тем более быстродействующей является система. В зависимости от соотношения параметров L_2 , L_3 , может быть различным тип расположения корней характеристического полинома замкнутой системы.

Разработаны различные методики выбора типа и параметров среднечастотного участка желаемых асимптотических ЛАЧХ, его «стыковки» с низкочастотным и высокочастотным участками, т.е. формирования желаемой передаточной функции разомкнутого контура $W_p(s)$.

Если ставится задача формирования свободных движений, то среднечастотная область желаемых ЛАЧХ формируется по методике, базирующейся на специальных диаграммах связи корней характеристического полинома замкнутой системы с нулями и полюсами передаточной функции разомкнутой системы [22]. На рис. 4.18 показана ЛАЧХ L_{p1}^* типа 2–1–3, параметры $L_2 = 6$ дБ, $L_3 = 17$ дБ которой выбраны по такой диаграмме из условия наибольшего быстродействия процессов — максимума степени устойчивости η , при допустимой колебательности корней $\mu \leq 1$.

Методика В. В. Солодовникова [75] предназначена для синтеза систем управления по требованиям к прямым показателям качества переходного процесса по каналу воспроизведения задающего воздействия. Исходными

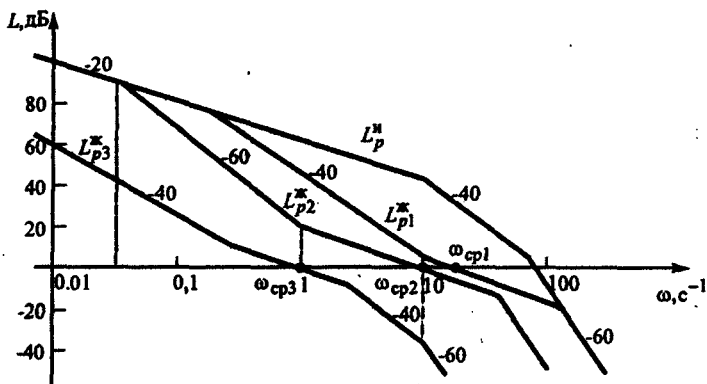


Рис. 4.18. Желаемые ЛАЧХ, сформированные по различным методикам

данными для формирования среднечастотной области желаемых ЛАЧХ являются время регулирования t_p и перерегулирование σ типовой переходной характеристики $h(t)$ — реакции системы на единичное ступенчатое задающее воздействие при нулевых предначальных условиях.

Предполагается, что замкнутая система имеет типовую вещественную частотную характеристику $P(\omega)$, аппроксимируемую тремя трапециями (рис. 4.19) с определенными соотношениями параметров. В частности, принимается, что $P_{\min} \approx 1 - P_{\max}$. При этом переходные процессы в системе имеют приемлемый вид, а их перерегулирование определяется в основном величиной $P_{\max}$. Номограмма В. В. Солодовникова [75] позволяет найти параметры $P_{\max}$ и ω_n по допустимым значениям σ и t_p . Далее по номограмме, связывающей частотные характеристики разомкнутой системы $L_p(\omega)$ и $\phi_p(\omega)$ с вещественной частотной характе-

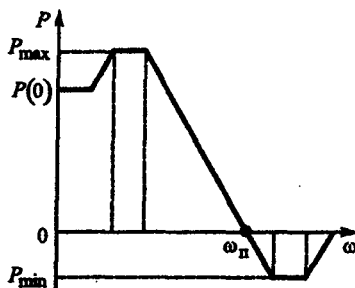


Рис. 4.19. Типовая вещественная частотная характеристика

ристической замкнутой системы $P(\omega)$, находят параметры $L_2 = L_3$. Частота среза $\omega_{\text{ср}}$ с частотой ω_n связана приближенным соотношением

$$\omega_{\text{ср}} \approx (0,6 \dots 0,9) \omega_n.$$

Желаемая асимптотическая ЛАЧХ $L_{\text{р2}}^*$, параметры среднечастотного участка которой выбраны по методике В. В. Солодовникова для значений $\sigma = 25\%$ и $t_p = 2$ с также показана на рис. 4.18.

Методика А. В. Фатеева [64] отличается от методики В. В. Солодовникова тем, что не требует равенства $L_2 = L_3$.

Методика В. А. Бесекерского [13] предназначена для формирования желаемых передаточных функций следящих систем, обрабатывающих эквивалентные гармонические воздействия с заданной установившейся ошибкой. Системы с астатизмом первого порядка должны иметь требуемое усиление на частоте воздействия, откуда находится необходимая добротность по скорости. В этой методике минимизируется полоса пропускания частот контура

$$\omega_{\text{ср}} \rightarrow \min$$

при ограничении на показатель колебательности

$$M = \max_{\omega} |\Phi(j\omega)|.$$

Параметры среднечастотного участка ЛАЧХ типа 2-1-2 выбираются из выражений

$$L_2 = 20 \lg \frac{M}{M-1};$$

$$L_3 = 20 \lg \frac{M}{M+1}.$$

На рис. 4.18 изображена асимптотическая желаемая ЛАЧХ $L_{\text{р3}}^*$ для исходных данных: $\omega_s = 0,0001 \text{ с}^{-1}$; $L_x = 110 \text{ дБ}$; $M = 1,3$. При этом $L_2 = 13 \text{ дБ}$; $L_3 = 5 \text{ дБ}$; $\omega_{\text{ср3}} = 1,0 \text{ с}^{-1}$, а желаемая передаточная функция имеет вид

$$W_{\text{р3}}^*(s) = \frac{1000(6,6s+1)}{s(10000s+1)(0,8s+1)(0,1s+1)}.$$

Методика А. А. Вавилова [64] предназначена для синтеза инвариантных до ϵ систем управления с высоким порядком астатизма. Предъявляются требования к переходным процессам переменной ошибки при степенных воздействиях. Выбор среднечастотных нулей и полюсов $W_{\text{р}}^*(s)$ осуществляется по специальным номограммам синтеза по критерию

минимального времени переходных процессов при заданном значении максимальной динамической ошибки.

В условиях применения программных средств для автоматизации синтеза систем управления роль приведенных выше номограмм сводится к выбору по ним начальных значений параметров. Далее в режиме оперативного взаимодействия с ЭВМ эти параметры уточняются.

4.8. ВЫЧИСЛЕНИЕ ПЕРЕДАТОЧНЫХ ФУНКЦИЙ КОРРЕКТИРУЮЩИХ УСТРОЙСТВ

Коррекция автоматической системы, по существу, означает очередное усложнение алгоритма управления для обеспечения устойчивости системы и требуемого качества переходных процессов. Усложнение алгоритма сводится к более глубокой переработке поступающей информации или связано с необходимостью получения дополнительной текущей информации.

На языке линейных математических моделей, представленных в форме структурных схем или графов, дополнение управляющих устройств корректирующими устройствами формулируется как введение новых звеньев или дуг с передаточными функциями $W^*(s)$, обеспечивающими приближение характеристик скорректированных систем $W^c(s)$ к желаемым $W^*(s)$.

Пусть модель корректируемой системы представлена в виде одно-контурной системы с единичной отрицательной обратной связью. Передаточная функция $W_0(s)$ описывает динамические свойства неизменяемой (силовой) части системы — собственно объекта управления с управляющим органом, исполнительного механизма, усилителей. Сюда же отнесем измерительные и преобразовательные элементы. Передаточная функция $W_y(s)$ характеризует информационно-алгоритмическую часть системы — управляющее устройство, выполняющее функции внутреннего компенсатора. Пусть оно рассчитано так, что удовлетворяются требования к установившимся процессам; однако переходные процессы расходятся или недопустимо колебательны.

Если для коррекции не привлекается дополнительная текущая информация, т. е. не предусматривается установка новых измерительных элементов и на объект не оказываются корректирующие воздействия по дополнительным каналам, то на структурных схемах звено коррекции может включаться последовательно с компенсатором W_y (рис. 4.20, а), параллельно с ним (рис. 4.20, б) или в виде обратной связи (рис. 4.20, в).

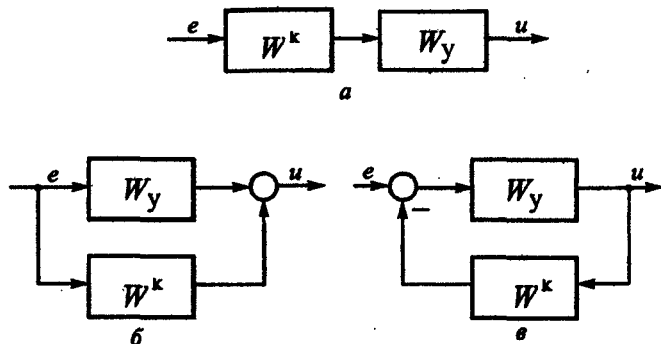


Рис. 4.20. Способы коррекции

Эти три способа в информационном и динамическом смыслах равносильны и обуславливают только особенности расчета передаточной функции $W^*(s)$, а также схемной (аппаратной) или программной реализации алгоритма коррекции и управляющего устройства в целом. Легко получить формулы взаимного пересчета передаточных функций для различных способов коррекции. Как будет показано далее, проще всего получить передаточную функцию последовательной коррекции.

В том случае, когда коррекция системы связана с установкой нового измерительного элемента, звено коррекции также образует контур местной обратной связи (рис. 4.21). Принципиальное отличие такой коррекции от рассмотренного выше случая (см. рис. 4.20, в) заключается в привлечении дополнительной текущей информации о некоторой внутренней переменной неизменяемой части системы. Корректирующее звено может охватывать усилитель, исполнительный механизм, часть объекта. На частотах, где усиление внутреннего контура велико, уменьшается чувствительность передаточной функции основного контура к вариациям оператора охваченной части, что было показано в § 3.9. Такой способ коррекции является наиболее часто рекомендуемым. Во многих электрических системах автоматического управления достаточно просто получить необходимую дополнительную текущую информацию.

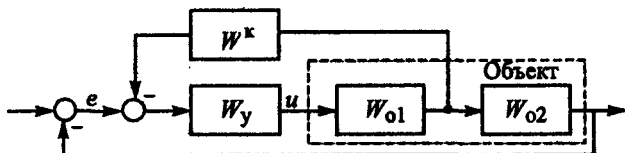


Рис. 4.21. Местная обратная связь

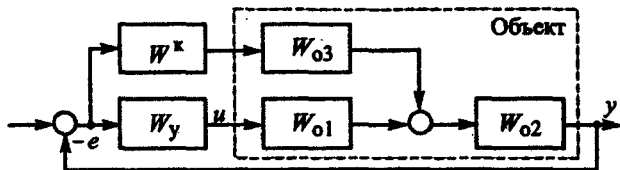


Рис. 4.22. Параллельная коррекция

Для оказания корректирующего воздействия на объект могут привлекаться новые каналы воздействий (рис. 4.22). Такой способ коррекции принципиально отличается от рассмотренного выше случая параллельной коррекции (см. рис. 4.20, б). Ввиду большей сложности изменения системы в энергетической (силовой) части такая коррекция применяется значительно реже.

Рассмотрим подробно процедуры вычисления передаточных функций последовательной коррекции и звена корректирующей обратной связи, а также некоторые свойства скорректированных систем.

4.8.1. Последовательная коррекция

Последовательная коррекция разомкнутого контура иллюстрируется на рис. 4.23, где передаточная функция исходного контура равна:

$$W^H(s) = W_0(s)W_y(s).$$

Передаточная функция скорректированного разомкнутого контура

$$W^c(s) = W^H(s)W^k(s) \quad (4.25)$$

должна быть тождественной желаемой передаточной функции

$$W^c(s) \equiv W^*(s). \quad (4.26)$$

Из тождества (4.26) с учетом (4.25) получим искомую передаточную функцию звена последовательной коррекции

$$W^k(s) = W^*(s) / W^H(s) = B^k(s) / A^k(s). \quad (4.27)$$

Это дробно-рациональная функция, полиномы числителя B^k и знаменателя A^k которой определяются так:



Рис. 4.23. Последовательная коррекция контура

$$\frac{B^*}{A^*} = \frac{B^* A^n}{A^* B^n}.$$

Если полиномы числителей и знаменателей передаточных функций W^n и W^* имеют общие делители

$$\text{НОД}(B^*, B^n) = B_2;$$

$$\text{НОД}(A^*, A^n) = A_2,$$

т. е.

$$B^* = B_1^* B_2; \quad B^n = B_1^n B_2; \quad A^* = A_1^* A_2; \quad A^n = A_1^n A_2, \quad (4.28)$$

то после сокращения общих делителей порядок передаточной функции звена коррекции понижается:

$$W^* = W_1^* = \frac{B_1^*}{A_1^*} = \frac{B_1^* A_1^n}{A_1^* B_1^n}.$$

При формировании желаемой передаточной функции W^* следует стремиться к повторению максимального числа нулей и полюсов W^n ; это упрощает коррекцию.

Пусть передаточная функция W_1^* реализована и введена в систему. Передаточная функция скорректированного разомкнутого контура равна:

$$W^c = \frac{B^c}{A^c} = \frac{B^n}{A^n} \frac{B_1^* A_1^n}{A_1^* B_1^n}.$$

С учетом соотношений (4.28) получим

$$W^c = W^* \frac{B_1^n A_1^n}{A_1^n B_1^n}.$$

Передаточная функция скорректированной системы равна желаемой с точностью до диполей, равных несовпадающим нулям и полюсам передаточных функций W^* и W^n . Передаточная функция звена коррекции имеет нули и полюсы, компенсирующие часть полюсов и нулей передаточной функции исходной системы. В результате передаточная функция скорректированного контура W^c оказывается *неполной*.

Характеристический полином скорректированной системы равен произведению желаемого характеристического полинома на компенсирующие друг друга полиномы

$$A^c = A^* A_1^n B_1^n.$$

Таким образом, при последовательной коррекции часть динамики как бы выводится из контура, так как на комплексных частотах компенсируемых полюсов усиление контура равняется нулю. Помимо этого в контур вводится желаемая динамика — желаемое соотношение полюсов и нулей.

Безусловно, полиномы B_1^n , A_1^n должны быть устойчивыми. Корни полиномов B_2 , A_2 могут быть любыми.

Порядок скорректированной системы

$$n^c = n^* + n_1^n + m_1^n$$

выше порядка желаемой системы на число полюсов и нулей неполной части. Комплексно-частотный метод дает желаемое поведение с точностью до искусственно сформированной неполной части системы.

Передаточные функции замкнутых скорректированных систем по воспроизведению

$$\Phi_g^c = \frac{W^c}{1+W^c} = \Phi^* \frac{B_1^n A_1^n}{A_1^n B_1^n}$$

и по ошибке

$$\Phi_e^c = \frac{1}{1+W^c} = \Phi_e^* \frac{B_1^n A_1^n}{A_1^n B_1^n}$$

имеют те же диполи, что и W^c . Поэтому, если среди диполов есть левые, но не очень хорошие, например, сильно колебательные, то на переходных процессах, определяемых передаточными функциями Φ_g^c и Φ_e^c , они не отражаются. Однако при произвольных предначальных условиях свободные движения могут определяться всеми корнями характеристического полинома.

Если к системе приложено возмущение, причем передаточная функция пути в графе от точки приложения до выхода системы имеет некомпенсированные неудовлетворительные полюсы, то они будут определять переходные процессы, так как не будут диполями соответствующей передаточной функции замкнутой системы. На рис. 4.24 изображена структурная схема, иллюстрирующая такую ситуацию. Передаточная функция скорректированной замкнутой системы по возмущению

$$\Phi_f^c = \frac{W_1^n}{1+W^c} = \Phi_f^* \frac{B_1^n}{A_1^n}$$

имеет полюсы, равные корням полинома A_1^n , но не имеет таких нулей. Переходные процессы системы при возмущениях будут определяться и

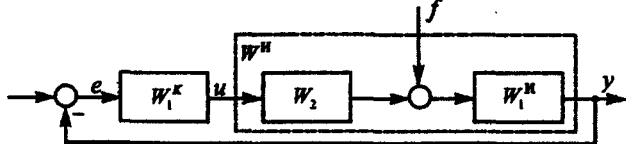


Рис. 4.24. Следящая система, к которой приложено возмущение

корнями λ^* , среди которых не должно быть сильно колебательных или недопустимо близких к мнимой оси.

Как следует из (4.27), ЛАЧХ звена последовательной коррекции равна разности желаемой и исходной ЛАЧХ

$$L^k(\omega) = L^*(\omega) - L^H(\omega).$$

ЛАЧХ контура равна сумме ЛАЧХ исходного контура и звена коррекции

$$L^c(\omega) = L^H(\omega) + L^k(\omega).$$

Она в точности равна желаемой ЛАЧХ L^* , так как частотные характеристики не отражают диполей. Поэтому в частотных методах, использующих графические или численные процедуры определения характеристик звеньев коррекции, следует заранее быть уверенным, что не будет плохих диполей передаточной функции W^c .

4.8.2. Местная обратная связь

Коррекция контура в виде местной обратной связи иллюстрируется рис. 4.25, где выделены охваченная и неохваченная части исходного контура

$$W^H = W^{ox} W^{no}.$$

Передаточная функция скорректированного разомкнутого контура

$$W^c = \frac{W^{ox}}{1 + W^{ox} W^k} W^{no} = \frac{W^H}{1 + W^{bk}}, \quad (4.29)$$

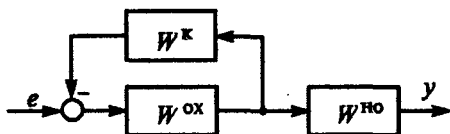


Рис. 4.25. Коррекция контура в виде местной обратной связи

где

$$W^{BK} = W^{OK} W^K$$

— передаточная функция внутреннего контура, должна быть тождественной желаемой передаточной функции.

Из тождества (4.26) с учетом (4.29) получим выражение для передаточной функции внутреннего контура

$$W^{BK} = \frac{W^H}{W^K} - 1$$

и искомой передаточной функции корректирующей обратной связи

$$W^K = \frac{W^{BK}}{W^{OK}} = \frac{1}{W^{OK}} \left(\frac{W^H}{W^K} - 1 \right).$$

Выражение для полиномов числителя и знаменателя передаточной функции внутреннего контура имеет вид

$$W^{BK} = \frac{B^{BK}}{A^{BK}} = \frac{B^H A^K - A^H B^K}{B^K A^H}.$$

Выражение для полиномов числителя и знаменателя передаточной функции звена коррекции зависят и от свойств охваченной части

$$W^K = \frac{B^K}{A^K} = \frac{B^H A^K - A^{HO} A^{OK} B_1^K}{B^K A^{HO} B^{OK}}.$$

Если полиномы числителя и знаменателя W^K имеют общие делители:

$$\text{НОД}(B^K, B^H) = B_2;$$

$$\text{НОД}(A^K, A^{HO}) = A_2,$$

то порядки передаточных функций W^K и W^{BK} понижаются:

$$W^K = W_1^K = \frac{B_1^K}{A_1^K} = \frac{B_1^H A_1^K - A_1^{HO} A^{OK} B_1^K}{B_1^K A_1^{HO} B^{OK}};$$

$$W^{BK} = W_1^{BK} = \frac{B_1^{BK}}{A_1^{BK}} = \frac{A_1^H A_1^K - A_1^{HO} A^{OK} B_1^K}{B_1^K A_1^{HO} A^{OK}}.$$

При формировании желаемой передаточной функции W^* целесообразно повторять нули W^H и полюсы W^{HO} , что упрощает коррекцию.

При выполнении условий

$$B_1^n(0)A_1^{\pi}(0) = A_1^{no}(0)A^{ox}(0)B_1^{\pi}(0)$$

или

$$W_1^{\pi}(0) = W_1^n(0)$$

свободный член полинома $B_1^{\pi}(s)$ равняется нулю. Здесь приняты обозначения:

$$W^{\pi} = W_1^{\pi} \frac{B_2}{A_2},$$

$$W^n = W_1^n \frac{B_2}{A_2}.$$

Пусть передаточная функция звена коррекции реализована в соответствии с выражением:

$$W_1^{\pi} \frac{B_1^{\pi}}{A_1^{\pi}} = \frac{\tilde{B}_1^n A_1^{\pi} - \tilde{A}_1^{no} \tilde{A}^{ox} B_1^{\pi}}{B_1^{\pi} \tilde{A}_1^{no} \tilde{B}^{ox}} \quad (4.30)$$

и введена в систему. Здесь волнистая линия (тильда) над обозначениями полиномов передаточных функций объекта обозначает приближенность их моделирования и реализации оператора звена коррекции. Полиномы числителя и знаменателя передаточной функции скорректированного контура в соответствии с формулами (4.29) и (4.30) равны:

$$\begin{aligned} W^c &= \frac{B_1^n B_1^{\pi} - \tilde{A}_1^{no} \tilde{A}^{ox}}{\left(A^{ox} B_1^{\pi} \tilde{A}_1^{no} \tilde{B}^{ox} + B^{ox} \tilde{B}_1^n A_1^{\pi} - \tilde{A}^{ox} B_1^{\pi} \tilde{A}_1^{no} B^{ox} \right) A^{no}} = \\ &= \frac{\tilde{B}^n B_1^{\pi} \tilde{A}_1^{no} \tilde{B}^{ox}}{B^{ox} \tilde{B}_1^n A_1^{\pi} A^{no}}. \end{aligned} \quad (4.31)$$

Примем во внимание соотношения (4.28) и запишем

$$W^c = \frac{B^{\pi} B_1^n \tilde{A}_1^{no} \tilde{B}^{ox}}{A^{\pi} \tilde{B}_1^n A_1^{no} B^{ox}}.$$

Оказалось, что скорректированная передаточная функция имеет диполи из числа нулей W^{ox} ; нулей W^n , которые не совпадают с нулями W^{π} ; полюсов передаточной функции W^{no} ; неохваченной части из числа тех, которые не совпадают с полюсами W^{π} . Все они ока-

жуются корнями характеристического полинома скорректированной замкнутой системы

$$A^c = A^* B_1^n A_1^{no} B^{ox}.$$

Как видно из (4.31), при получении полиномов передаточной функции скорректированной системы W^c происходит компенсация слагаемых в знаменателе. Условием компенсации является совпадение модели охваченной части с оригиналом

$$\frac{\tilde{B}^{ox}}{\tilde{A}^{ox}} = \frac{B^{ox}}{A^{ox}}.$$

Поскольку точного равенства нет, нет и полной компенсации слагаемых. Поэтому необходимо установить, к чему может привести неточная компенсация.

Компенсация слагаемых не должна приводить к понижению степени полинома знаменателя $W^c(s)$, т. е. порядка скорректированной системы, так как в противном случае она будет негрубой — малые вариации параметров приведут к потере устойчивости. Система будет грубой, если взаимное уничтожение слагаемых не приведет к понижению степени полинома. Это имеет место, если степень остающегося слагаемого не ниже степеней уничтожающихся

$$\deg(B^{ox} \tilde{B}_1^n A_1^*) \geq \deg(A^{ox} B_1^* \tilde{A}_1^{no} \tilde{B}^{ox}),$$

т. е.

$$m^{ox} + m_1^n + n_1^* \geq n^{ox} + m_1^* + n_1^{no} + m^{ox}.$$

Прибавим к обеим частям $m_2 + n_2$, где $m_2 = \deg B_2$; $n_2 = \deg A_2$, и получим

$$n^* + m^n \geq n + m.$$

Условие грубости скорректированной системы примет вид

$$n^* - m^* \geq n^n - m^n.$$

Таким образом, разность степеней полиномов знаменателя и числителя желаемой передаточной функции W^* должна быть не ниже соответствующей разности для исходной передаточной функции W^n .

Как следует из (4.29), для значений аргумента s , удовлетворяющих условию $|W^*(s)| \gg 1$, справедливо приближенное выражение

$$W^c(s) \approx W^{no}(s) / W^*(s).$$

Это область комплексных частот, где усиление контура коррекции велико. Именно на этих частотах существенно влияние корректирующей

обратной связи. Отсюда получается передаточная функция звена коррекции

$$W^k(s) \approx W^{no}(s)/W^*(s),$$

приближенно обеспечивающая тождество (4.26). Если для минимально-фазовых передаточных функций оперировать ЛАЧХ, то для существенных частот коррекции имеет место:

$$L^k(\omega) \approx L^{no}(\omega) - L^*(\omega).$$

Для значений аргумента s , удовлетворяющих условию

$$|W^{*k}(s)| \ll 1,$$

справедливо другое приближенное выражение

$$W^c(s) \approx W^h(s).$$

На этих комплексных частотах контур коррекции практически не влияет на свойства системы.

4.9. ПАРАМЕТРИЧЕСКИЙ СИНТЕЗ СИСТЕМ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Синтез называют параметрическим, если заданы топология системы, структуры операторов звеньев, конкретизирована часть параметров и установлены пределы изменения для остальных. Задача параметрического синтеза часто сводится к расчету настроек регуляторов промышленной автоматики, когда закон регулирования уже выбран. Решение и других задач синтеза систем автоматического управления — обеспечения инвариантности к возмущениям, стабилизации объектов — завершается этапом параметрического синтеза. Этот этап характеризуется относительно малой исходной неопределенностью — на предыдущих этапах проектирования приняты принципиальные решения о точках съема текущей информации и приложения воздействий на объект, выбраны алгоритмы переработки информации, конкретизирована часть параметров и установлены допустимые пределы для других.

К системам автоматического управления предъявляется совокупность требований — инвариантность к воздействиям, устойчивость движений, малая чувствительность к вариациям операторов. Установившиеся и переходные, основные и дополнительные процессы количественно характеризуются показателями качества. При параметричес-

ком синтезе преобладает тенденция к взвешенному учету всей совокупности требований к поведению систем автоматического управления.

В постановках задач параметрического синтеза требования к процессам управления могут выражаться в виде ограничений на показатели качества

$$I_i(q_1, q_2, \dots) \leq I_{i\max}; \quad i = 1, 2, \dots \quad (4.32)$$

Параметры системы должны принадлежать допустимой области Q , например, заданным интервалам:

$$q_{k\min} \leq q_k \leq q_{k\max}; \quad k = 1, 2, \dots$$

В результате решения задачи должно быть найдено подмножество значений параметров $Q^* \subseteq Q$, при которых удовлетворяются все ограничения (4.32).

Задача параметрического синтеза в такой постановке может иметь множество решений, единственное решение или не иметь ни одного. На рис. 4.26, а, б иллюстрируются ситуации: 1 — множество решений; 2 — отсутствие решения — для случая, когда выбираются два параметра q_1 и q_2 и предъявлены требования к двум показателям качества $I_1(q_1, q_2)$, $I_2(q_1, q_2)$. На плоскости показателей качества (I_1, I_2) отображается область допустимых параметров Q .

Отсутствие пересечения области допустимых параметров с областью требуемых значений показателей качества свидетельствует об отсутствии решения. Это обычно означает, что этапы топологического и структурного синтеза проведены неудовлетворительно. Следует пересмотреть ограничения на параметры, требования к показателям качества, пересмотреть структуру операторов и даже топологию системы,

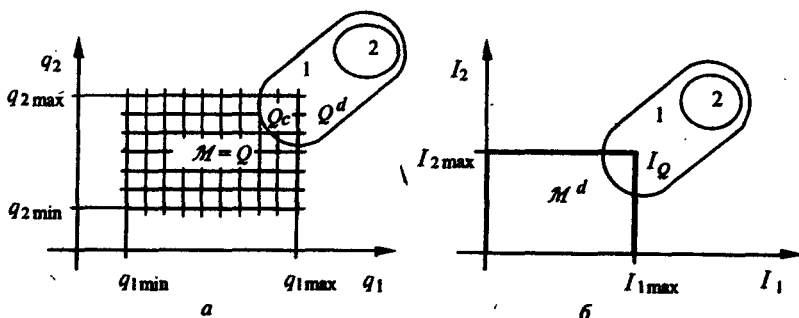


Рис. 4.26. Иллюстрация задачи параметрического синтеза

что изменяет зависимости $I_i(q_1, q_2, \dots)$, увеличивает число выбираемых параметров.

Отображение области допустимых значений параметров Q на пространство показателей качества в условиях применения развитых программных средств не представляет больших трудностей, если невелики размерности пространств. Это задача многократного анализа, правда, связанная с необходимостью выбора конечного подмножества элементов Q . Как правило, выбирают узлы решетки, равноотстоящие в обычном (для малых интервалов) или логарифмическом масштабе. Более сложно решить обратную задачу (синтез) отображения области допустимых значений показателей качества на пространство параметров. Имеется ряд методов и программ, позволяющих строить в пространстве параметров поверхности равных значений показателей качества [35]. В условиях применения машинной графики эти подходы играют все более значительную роль.

Не всегда известны ограничения на показатели качества, или их нецелесообразно назначать. Тогда задача параметрического синтеза часто имеет оптимизационную постановку — некоторые показатели качества требуется, например, минимизировать

$$I_i(q_1, q_2, \dots) \rightarrow \min_{q \in Q^*}; \quad i = 1, 2, \dots \quad (4.33)$$

при ограничениях на другие. В (4.33) положено, что вектор параметров q принадлежит подмножеству Q^* , где соблюдены ограничения на другие показатели качества. В частности, непустое множество должно соответствовать устойчивым системам.

Если оптимизируется единственный показатель качества, то имеет место задача поиска условного экстремума функции нескольких переменных, которая решается методами нелинейного программирования.

Оптимизация систем управления, как правило, ведется по нескольким критериям (4.33). Имеет место задача *векторной оптимизации*

$$I(q) \rightarrow \min_{q \in Q}.$$

Компоненты I_i векторного критерия I называются частными критериями (показателями). Результаты минимизации частных критериев в общем случае не совпадают. На рис. 4.27 иллюстрируется задача минимизации двух частных критериев $I_1(q)$, $I_2(q)$, зависящих от одного параметра. Поскольку оптимальные по отдельным критериям значения параметров не совпадают, т.е. $q_1^* \neq q_2^*$, то возникает вопрос, что следует считать решением задачи векторной оптимизации.

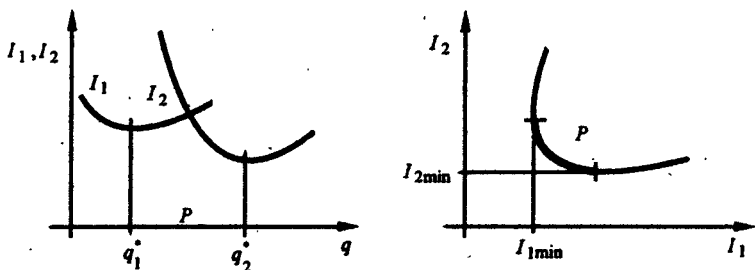


Рис. 4.27. Иллюстрация задачи векторной оптимизации

По определению, решением задачи векторной оптимизации является множество значений параметров, в котором изменение любого параметра с целью улучшения одного из частных критериев обязательно ухудшает хотя бы один другой. Такое множество неулучшаемых решений называют *множеством Парето* (V. Pareto).

На рис. 4.27, а области Парето P на оси параметра q соответствует отрезок между значениями q_1^* и q_2^* частных оптимумов. Здесь изменение q с целью уменьшения одного показателя приводит к увеличению другого. Иная иллюстрация области Парето приведена на рис. 4.27, б, где по осям отложены значения частных показателей $I_1(q)$ и $I_2(q)$.

Неполная определенность решения задачи векторной оптимизации (множественный характер решения) обусловлена неопределенностью постановки задачи. При формализации пожеланий проектировщика не установлены предпочтения и приоритеты. Уменьшение неопределенности решения связано с привлечением дополнительной информации.

Можно выделить два способа скаляризации задачи векторной оптимизации:

- все частные критерии, кроме одного, объявляются ограничениями;
- из частных критериев качества формируется скалярный критерий, который называют *обобщенным* или *глобальным*:

$$J(q, c) \rightarrow \min, \quad q \in Q$$

и его оптимум зависит от ряда весовых коэффициентов $c_i \geq 0; i = 1, 2, \dots$, при назначении которых предпочтение отдают тем или иным частным показателям. Обобщенный критерий называют *эффективным*, если для любых значений весовых коэффициентов его минимум принадлежит области Парето. Меняя весовые коэффициенты и решая задачу скалярной оптимизации методами нелинейного программирования, находят точки области Парето.

Примером эффективного критерия является аддитивный обобщенный критерий:

$$J(\mathbf{q}, \mathbf{c}) = \sum_i c_i I_i(\mathbf{q}); \quad \sum_i c_i = 1.$$

Рассмотренная в п. 3.11.2 улучшенная интегральная оценка является примером аддитивного обобщенного показателя с весовыми коэффициентами, равными единице и τ_1^2 . Частными показателями являются интегральные квадратичные оценки переходной составляющей процесса и его производной по времени. Там же приведено выражение для аддитивной обобщенной интегральной оценки по переменным состояниям, весовые коэффициенты которых упорядочены в матрицу. В п. 4.3.3 рассмотрен аналитический синтез оптимальных регуляторов по обобщенному критерию (4.13), являющийся примером параметрического синтеза.

Пример 5. Рассмотрим задачу параметрического синтеза следящей системы привода руля судна (см. рис. 1.2) по векторному критерию. Пусть выбрана топология системы. На рис. 4.28 изображен граф ее упрощенной линейной модели. Задачу структуры операторов дуг и часть параметров. Имеем астатическую систему, на вход которой подается воздействие $g(t) = g_1 t$.

Необходимо выбрать значение коэффициента передачи разомкнутого контура — добротности по скорости

$$k_1 > 0.$$

Пусть векторный критерий качества состоит из двух частных критериев:

$$I_1(k_1) = e_{\text{уст}}(k_1) \rightarrow \min;$$

$$I_2(k_1) = \mu(k_1) \rightarrow \min,$$

где установившаяся ошибка обратно пропорциональна добротности

$$e_{\text{уст}} = q_1 / k_1,$$

а $\mu(k_1)$ — колебательность корней характеристического полинома замкнутой системы

$$A_3(s) = Ts^2 + s + k_1.$$

Колебательность корней — отношение модулей мнимых частей корней к модулю действительной части (см. пр. 2 из п. 3.11.2):

$$\mu(k_1) = \begin{cases} 0, & k_1 \leq 1/4T; \\ \sqrt{4Tk_1 - 1}, & k_1 > 1/4T. \end{cases}$$

Система при $k_1 > 0$ всегда устойчива. Но для больших значений $k_1 > 1/4T$ комплексные корни $A_3(s)$ могут быть сильно колебательными.

На рис. 4.29 изображены графики зависимости частных показателей качества от добротности системы.

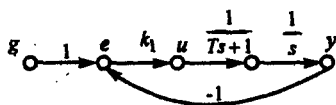


Рис. 4.28. Граф системы к примеру параметрического синтеза

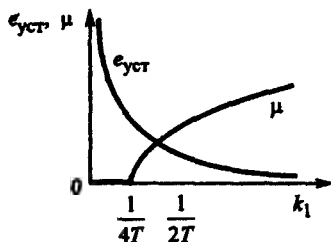


Рис. 4.29. Графики зависимости показателей качества от добротности контура

Очевидно, отрезок $(0, 1/4T]$ на оси абсцисс не является областью Парето, так как здесь увеличение k_1 уменьшает установившуюся ошибку, не увеличивая колебательности процессов. Областью Парето, т. е. решением задачи векторной оптимизации, являются значения $k_1 > 1/4T$. В этой области любое изменение k_1 , приводя к уменьшению одного критерия, обязательно увеличивает другой.

Таким образом, формализация требований к системе управления в виде векторного критерия позволила несколько сократить исходную неопределенность относительно добротности системы — вместо $k_1 > 0$ имеем $k_1 > 1/4T$. Дальнейшее уменьшение неопределенности решения связано с дополнительной информацией о предпочтениях проектировщика, позволяющей достичь компромисса между интересами отдельных частных критериев.

ЧАСТЬ 3

ТЕОРИЯ ЛИНЕЙНЫХ ДИСКРЕТНЫХ СИСТЕМ УПРАВЛЕНИЯ

Глава 5

ОБЩИЕ СВЕДЕНИЯ О ДИСКРЕТНЫХ СИСТЕМАХ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

5.1. КЛАССИФИКАЦИЯ ДИСКРЕТНЫХ СИСТЕМ ПО ВИДУ КВАНТОВАНИЯ

В рассмотренных выше непрерывных системах автоматического управления рабочая информация была представлена в виде сигналов, описываемых непрерывными функциями времени. Наряду с непрерывными способами передачи и преобразования сигналов широко используются дискретные, в которых осуществляется квантование сигналов. Квантование или дискретизация состоит в представлении непрерывного сигнала его дискретными значениями. В зависимости от вида квантования системы автоматического управления подразделяют на дискретные по уровню (релейные), дискретные по времени (импульсные), дискретные по уровню и по времени (релейно-импульсные) [88].

В *релейных системах* производится квантование по уровню δ_0 (рис. 5.1), при котором в произвольные моменты времени выделя-

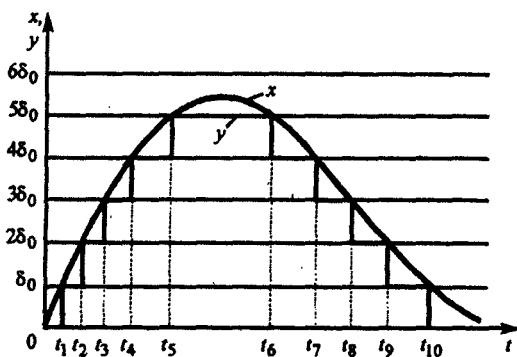


Рис. 5.1. Квантование по уровню

ются значения непрерывного сигнала при достижении им фиксированных дискретных уровней.

В импульсных системах производится квантование по времени, при котором в дискретные моменты времени выделяются значения непрерывного сигнала (рис. 5.2). Обычно эти моменты времени отстоят друг от друга на некоторую постоянную величину T , называемую *периодом квантования*. Квантование по времени осуществляется импульсным элементом, выходной сигнал которого представляет собой последовательность импульсов.

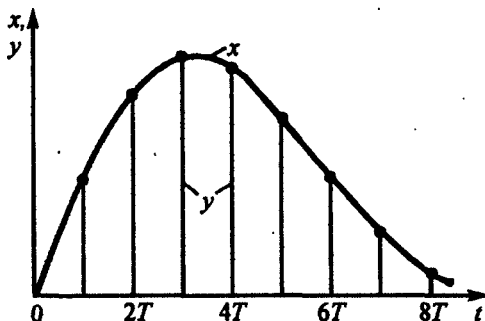


Рис. 5.2. Квантование по времени

В релейно-импульсных системах производится квантование по уровню и по времени, при котором непрерывный сигнал заменяется фиксированными дискретными (по уровню) значениями, ближайшими к значениям x в дискретные моменты времени (рис. 5.3). Квантование по уровню и по времени осуществляется в цифровых системах управле-

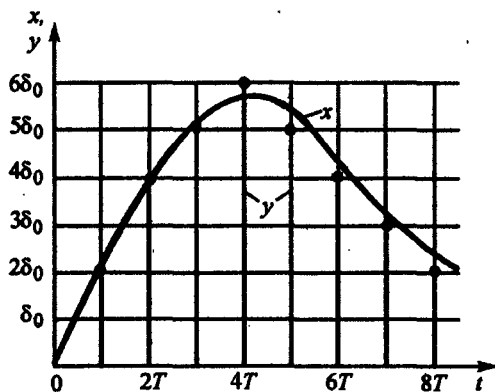


Рис. 5.3. Квантование по уровню и по времени

ния при преобразовании непрерывных сигналов в цифровую форму или в коды, поэтому в общем случае цифровые системы управления относятся к классу релейно-импульсных систем [56].

Область применения дискретных систем в настоящее время очень разнообразна. Можно выделить две основные категории таких систем.

- Первая — системы, которые дискретны по своей физической природе, т.е. информация в них существует только в дискретные моменты времени. Примерами этой категории дискретных систем являются радиолокационные системы обнаружения и сопровождения цели. Существуют многочисленные физические и биологические явления, процессы в социальных и экономических сферах, динамика которых адекватно может быть описана только дискретными моделями.

- Вторая категория — системы, в которых информация существует непрерывно, но намеренно квантуется для получения некоторых новых свойств по сравнению с непрерывными системами. Такими свойствами могут быть: простота реализации, повышение надежности, увеличение точности, меньшие габаритные размеры и стоимость. Быстрый прогресс в вычислительной технике, широкое использование микропроцессоров в системах управления еще более увеличивают интерес к дискретным системам.

5.2. ПОНЯТИЯ ОБ ИМПУЛЬСНЫХ СИСТЕМАХ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Импульсная система отличается от непрерывной наличием импульсного элемента, осуществляющего квантование по времени. Процесс преобразования непрерывного сигнала в последовательность импульсов, параметры которых зависят от значений этого сигнала в дискретные моменты времени, называют *импульсной модуляцией*. Входным сигналом импульсного элемента или модулятора является непрерывный сигнал, а выходным — модулированная последовательность импульсов.

В зависимости от того, какой параметр импульса (амплитуда, длительность, фаза) модулируется непрерывным сигналом (рис. 5.4), различают: амплитудно-импульсную модуляцию (АИМ) — рис. 5.4, а, широтно-импульсную модуляцию (ШИМ) — рис. 5.4, б, фазоимпульсную модуляцию (ФИМ) — рис. 5.4, в. Возможна также модуляция, при которой амплитуда, длительность и фаза импульсов постоянны, а функцией непрерывного сигнала на входе модулятора является период повторения или частота импульсов на выходе модулятора. Такой вид модуляции называется частотно-импульсной (ЧИМ). Частным случаем

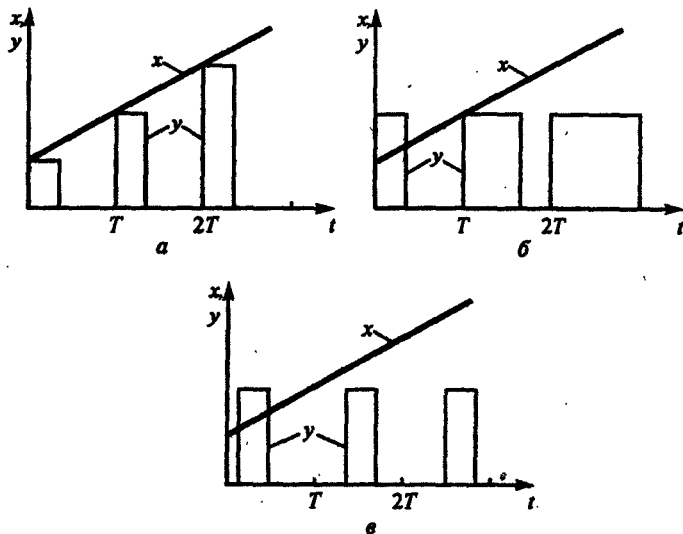


Рис. 5.4. Виды импульсной модуляции

ЧИМ является числоимпульсная модуляция, при которой изменяется число импульсов на определенном временном отрезке.

Если модулируемый параметр последовательности импульсов определяется значениями входного сигнала в фиксированные равноотстоящие моменты времени и остается постоянным в течение времени существования импульса, то такой вид модуляции называется импульсной модуляцией *первого рода*. Возможны случаи, когда модулируемый параметр последовательности импульсов в течение времени существования импульса изменяется в соответствии с текущим значением входного сигнала. Такой вид модуляции называется импульсной модуляцией *второго рода*. Примером систем с амплитудно-импульсной модуляцией второго рода являются системы прерывистого регулирования или системы с конечным временем замыкания импульсного элемента (рис. 5.5).

Системы автоматического управления в зависимости от вида импульсной модуляции подразделяются на амплитудно-импульсные системы (АИС), широтно-импульсные системы (ШИС), фазоимпульсные системы (ФИС) и т.д. Амплитудно-им-

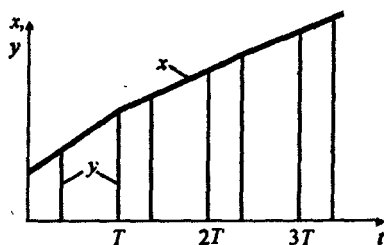


Рис. 5.5. Модуляция второго рода

пульсные системы могут быть как линейными, так и нелинейными (при наличии нелинейного элемента). Широтно-импульсные системы принципиально нелинейны. Нелинейными являются и цифровые системы, в которых имеет место квантование по уровню.

Зависимость модулируемого параметра M последовательности импульсов x_M на выходе импульсного элемента от соответствующих значений входного сигнала x называется статической характеристикой импульсного элемента. Коэффициентом передачи (усиления) импульсного элемента является отношение значения модулируемого параметра M выходной последовательности импульсов к значению входного сигнала x . В общем случае характеристика импульсного элемента может быть линейной и нелинейной. На линейном участке крутизна характеристики импульсного элемента равна значению его коэффициента усиления.

На рис. 5.6 показана статическая характеристика амплитудно-импульсного элемента. Коэффициент передачи такого элемента в области малых значений x равен

$$k_x = \frac{A}{x} = \operatorname{tg} \alpha,$$

где модулируемый параметр M — амплитуда импульса A . Аналогичная характеристика для широтно-импульсного элемента показана на рис. 5.7. Коэффициент передачи широтно-импульсного элемента

$$k_x = \frac{T_n}{x} = \frac{\gamma T}{x},$$

где модулируемый параметр M — длительность импульса T_n , а $\gamma = \frac{T_n}{T}$ — относительная длительность импульса.

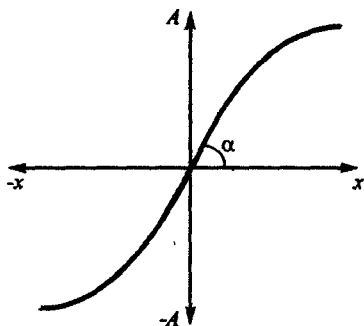


Рис. 5.6. Статическая характеристика АИМ

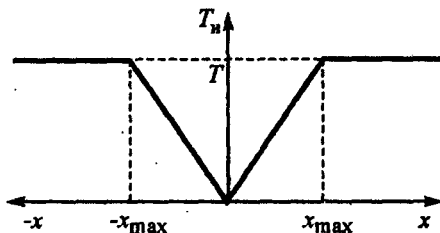


Рис. 5.7. Статическая характеристика ШИМ

5.3. ОБОБЩЕННЫЕ СТРУКТУРНЫЕ СХЕМЫ ИМПУЛЬСНЫХ АВТОМАТИЧЕСКИХ СИСТЕМ

В общем случае любая однокоитурная импульсная автоматическая система (ИАС) может быть представлена взаимодействующими друг с другом импульсной ИЧ и непрерывной НЧ частями (рис. 5.8). В непрерывную часть обычно входит объект управления, а также усилитель мощности и исполнительное устройство. Импульсная часть, как правило, является управляющим устройством и объединяет функциональные элементы, участвующие в импульсном преобразовании сигнала. Эта часть может быть реализована в виде ключей, модуляторов, импульсных регуляторов, цифровых вычислительных устройств с аналого-цифровыми и цифроаналоговыми преобразователями и т.д.

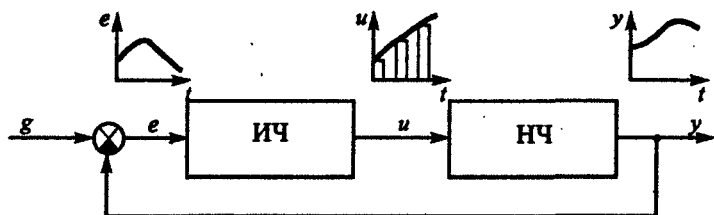


Рис. 5.8. Обобщенная структурная схема замкнутой импульсной системы

Функционально импульсную часть можно рассматривать как некоторый преобразователь непрерывного сигнала в импульсное управляющее воздействие того или иного вида. В линейных амплитудно-импульсных системах выходной сигнал импульсной части представляет собой последовательность импульсов, амплитуды которых пропорциональны значениям непрерывного сигнала ошибки в равноотстоящие моменты квантования. В простейшем случае импульсная часть является реальным импульсным элементом (РИЭ) или импульсным модулятором (рис. 5.9).

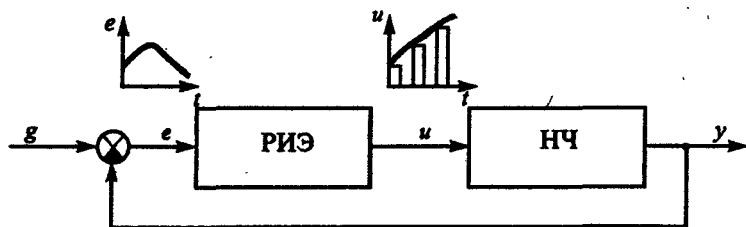


Рис. 5.9. Структурная схема импульсной системы с реальным импульсным элементом

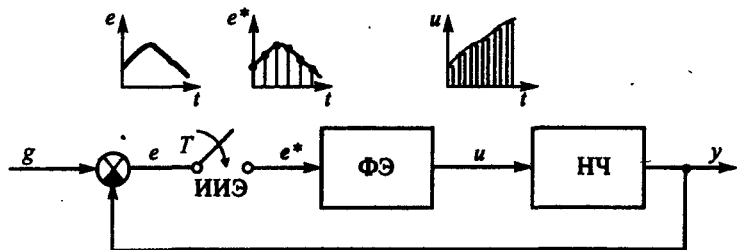


Рис. 5.10. Структурная схема импульсной системы с идеальным импульсным элементом

При исследовании импульсных систем их реальные импульсные элементы обычно заменяют последовательными соединениями идеального импульсного и формирующего элементов. Идеальный импульсный элемент (ИИЭ) под воздействием непрерывного входного сигнала формирует идеальные мгновенные импульсы вида δ -функций, «амплитуды площадей» которых равны значениям входного сигнала в моменты квантования (рис. 5.10). Формирующий элемент (ФЭ) преобразует эти импульсы в импульсы, соответствующие выходной импульсной последовательности $u(t)$ реального импульсного элемента (см. рис.5.9). Идеальными импульсами будем считать импульсы бесконечно большой амплитуды и бесконечно малой длительности с площадями, численно равными значениям непрерывного сигнала $e(t)$ в дискретные моменты квантования. Другими словами, ИИЭ представляет собой амплитудно-импульсный модулятор, непрерывный входной сигнал которого $e(t)$ модулирует несущую последовательность мгновенных δ -импульсов бесконечно большой амплитуды с единичной площадью, т.е. импульсов вида δ -функций с постоянным периодом повторения T , но не единичной, а переменной площади, которая является мерой входного сигнала в дискретные моменты времени. Формирующий элемент представляет собой некоторый непрерывный фильтр, реакция которого на мгновенный импульс последовательности $e^*(t)$ совпадает по своей форме с реальным импульсом последовательности $u(t)$ на выходе реального импульсного элемента (см. рис.5.9). Эта реакция при площади мгновенного импульса, равной единице, представляет собой импульсную (весовую) характеристику формирующего элемента $w_{\Phi}(t)$. Так, например, для единичного импульса прямоугольной формы (рис. 5.11) на интервале $0 \leq t < T$ имеем

$$u(t) = 1(t) - 1(t - \gamma T) = w_{\Phi}(t)$$

или

$$w_{\Phi\Xi}(t) = u(t) = \begin{cases} 1(t) & \text{при } 0 \leq t < \gamma T; \\ 0 & \text{при } \gamma T \leq t < T. \end{cases}$$

Поскольку изображение по Лапласу импульсной (весовой) характеристики звена является его передаточной функцией, передаточная функция ФЭ равна

$$H(s) = L\{u(t)\},$$

т.е. для прямоугольного импульса имеем

$$H(s) = \frac{1 - e^{-\gamma s T}}{s}.$$

При $\gamma = 1$ передаточная функция такого ФЭ, который называют экстраполятором нулевого порядка или фиксатором равна

$$H(s) = \frac{1 - e^{-sT}}{s}. \quad (5.1)$$

Фиксатор запоминает площадь мгновенного импульса на входе до прихода следующего импульса (рис. 5.12).

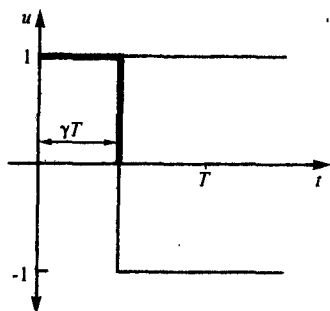


Рис. 5.11. Формирование прямоугольного импульса из двух ступенчатых функций

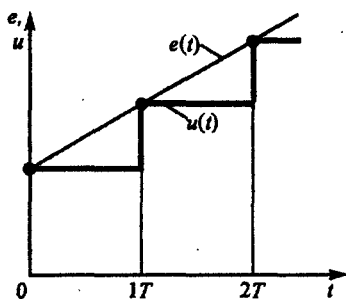
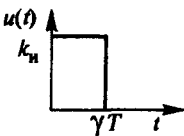
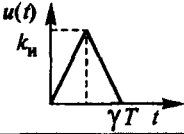
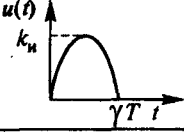


Рис. 5.12. Амплитудно-импульсная модуляция с фиксатором нулевого порядка

В табл. 5.1 приведены передаточные функции ФЭ для некоторых типовых форм импульсов.

Для удобства анализа ФЭ и непрерывную часть импульсной системы объединяют. В этом случае независимо от формы реальных импульсов импульсные системы (типа показанной на рис. 5.10) с амплитудной модуляцией первого рода можно представить в виде соединения идеального импульсного элемента и приведенной непрерывной части. При-

Форма импульса	$H(s)$	$H(s)$ при $\gamma \ll 1$
	$k_n \frac{1 - e^{-\gamma Ts}}{s}$	$k_n \gamma T$
	$k_n \frac{2(1 - e^{-\gamma Ts/2})^2}{\gamma Ts^2}$	$k_n \frac{\gamma T}{2}$
	$\frac{2k_n T \frac{\pi}{\gamma}}{T^2 s^2 (\frac{\pi}{\gamma})^2} (1 + e^{-\gamma Ts})$	$\frac{2k_n T \gamma}{\pi}$

веденной непрерывной частью (ПНЧ) называют последовательное соединение формирующего элемента и непрерывной части (рис.5.13). Передаточная функция приведенной непрерывной части равна

$$W_{\text{пнч}}(s) = H(s)W_{\text{нч}}(s).$$

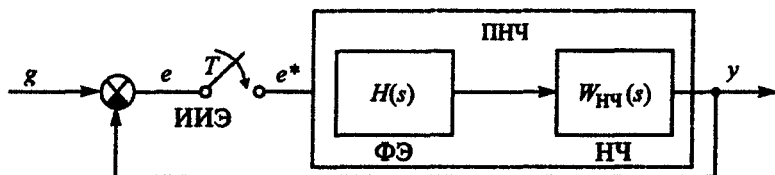


Рис. 5.13. Эквивалентная схема импульсной системы с амплитудно-импульсной модуляцией

Обычно коэффициент усиления импульсного элемента k_n относят к непрерывной части системы, считая, что коэффициент передачи формирующего элемента равен единице.

В цифровых системах управления (рис. 5.14) непрерывные сигналы квантуются по времени и по уровню. Квантование по времени осуществляется мультиплексором с периодом дискретности T , а квантование по уровню (округление или усечение) — в аналого-цифровом преобразователе (АЦП). Мультиплексор обычно конструктивно объединяется с устройством масштабирования и аналого-цифровым преобразователем. Цифровой сигнал обрабатывается процессором (цифровым вычислителем — ЦВ) и поступает в цифроаналоговый преобразователь

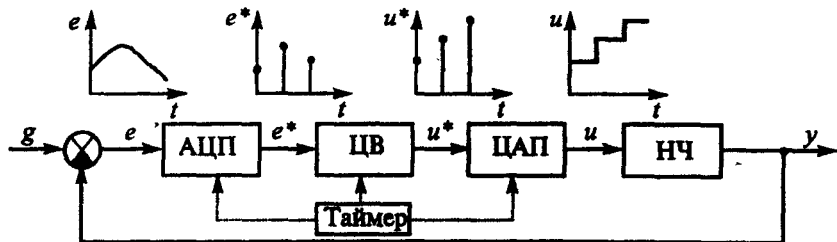


Рис. 5.14. Блок-схема цифровой системы управления

(ЦАП). Время τ_{κ} , затраченное на преобразование аналогового сигнала $e(t)$ в цифровую форму и последующую его обработку в процессоре, представляет собой задержку в канале управления. Обычно τ_{κ} значительно меньше постоянных времени исполнительного устройства и объекта, поэтому величиной задержки τ_{κ} часто пренебрегают, полагая, что входной и выходной квантователи работают синхронно. Таким образом, обобщенная структурная схема цифровой автоматической системы может быть представлена в виде, приведенном на рис. 5.15. Идеальные импульсные элементы соответствуют квантователям на входе и выходе процессора, сам процессор представлен в виде дискретного фильтра (ДФ), а квантование по уровню учитывается многоступенчатыми нелинейными элементами (НЭ).

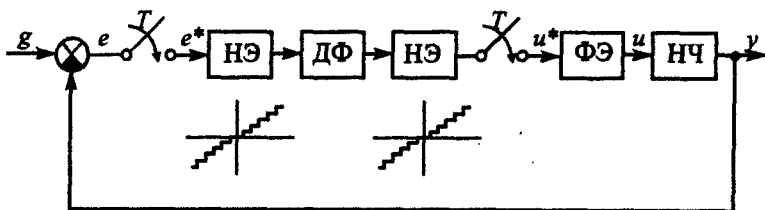


Рис. 5.15. Обобщенная структурная схема цифровой системы управления

При использовании аналого-цифровых и цифроаналоговых преобразователей, имеющих достаточно большое число разрядов, эффекты квантования по уровню часто можно не учитывать. Структурная схема цифровой системы управления в этом случае соответствует амплитудно-импульсной системе с импульсной частью в виде линейного дискретного фильтра (ЛДФ) (рис. 5.16). Выходной сигнал ЛДФ преобразуется формирующим устройством типа экстраполятора нулевого порядка (фиксатора).

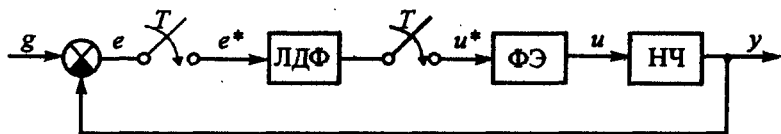


Рис. 5.16. Структурная схема цифровой системы управления без учета квантования по уровню

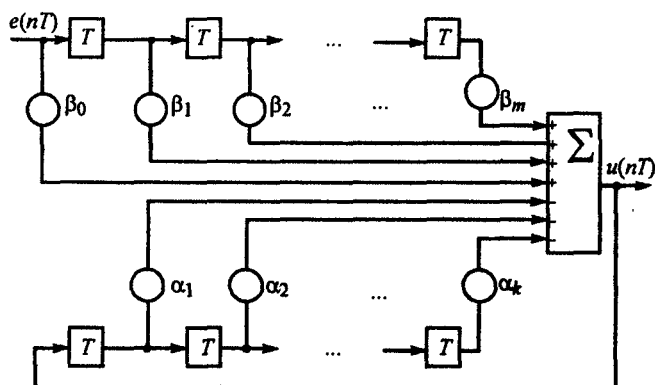


Рис. 5.17. Блок-схема линейного дискретного фильтра

В общем случае линейный дискретный фильтр осуществляет преобразование входной последовательности данных $\{e(iT)\}$ в выходную $\{u(iT)\}$; $i = 0, 1, 2, \dots$ в соответствии с уравнением

$$u(nT) = \sum_{i=0}^m \beta_i e[(n-i)T] - \sum_{i=1}^k \alpha_i u[(n-i)T].$$

Такое преобразование можно представить в виде блок-схемы, показанной на рис. 5.17, где T — блок памяти.

Глава 6

МОДЕЛИ ЛИНЕЙНЫХ ДИСКРЕТНЫХ СИСТЕМ УПРАВЛЕНИЯ

6.1. РЕШЕТЧАТЫЕ ФУНКЦИИ И РАЗНОСТНЫЕ УРАВНЕНИЯ

Как отмечалось, приведенная непрерывная часть реагирует лишь на дискретные значения непрерывного сигнала в моменты квантования nT . Поэтому непрерывную функцию $x(t)$, описывающую непрерыв-

ный сигнал, можно заменить соответствующей решетчатой функцией $x(nT)$, причем

$$x(nT) = x(t) \text{ при } t = nT;$$

$$x(nT) = 0 \text{ при } nT < t < (n+1)T,$$

где $n = 0, 1, 2, \dots$

Таким образом, для того чтобы получить решетчатую функцию по заданной непрерывной функции $x(t)$, нужно в последней заменить t на nT (рис. 6.1).

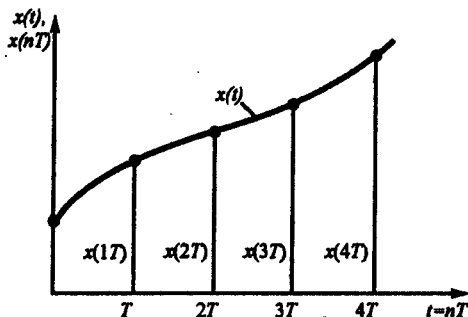


Рис. 6.1. Непрерывная функция $x(t)$ и соответствующая ей решетчатая функция $x(nT)$

Например, непрерывной функции $x(t) = e^{-\alpha t}$ соответствует решетчатая функция

$$x(nT) = x(t) \Big|_{t=nT} = e^{-\alpha nT},$$

а функции $x(t) = kt$ соответствует решетчатая функция

$$x(nT) = x(t) \Big|_{t=nT} = knT.$$

Решетчатые функции описывают «порождающие» их непрерывные функции только в дискретные моменты времени, совпадающие с моментами квантования. В промежутках между моментами квантования информация об изменениях непрерывных функций отсутствует. Если интервал квантования T задан, то по функции $x(t)$ решетчатая функция $x(nT)$ определяется однозначно. Обратное утверждение несправедливо. Очевидно, что различным непрерывным функциям $x_i(t)$ может соответствовать одна и та же решетчатая функция $x(nT)$ (рис. 6.2). Поэтому по решетчатой функции $x(nT)$ невозможно определить «породившую» ее непрерывную функцию.

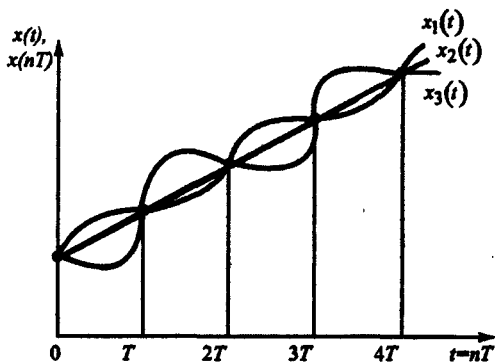


Рис. 6.2. Решетчатая функция $x(nT)$ и порождающие ее непрерывные функции

Для выявления поведения непрерывной функции между моментами квантования вводят промежуточное фиксированное время $\Delta t = \sigma T$ (рис. 6.3), которое изменяется от 0 до T . В этом случае непрерывную функцию $x(t)$ можно заменить смещенной решетчатой функцией

$$x(nT, \sigma T) = x(t) \text{ при } t = nT + \sigma T,$$

где n — любое целое число, а $0 \leq \sigma T \leq T$. Изменяя σT от 0 до T , можно получить семейство решетчатых функций $x(nT, \sigma T)$; $n = 1, 2, \dots$, определяющее функцию $x(t)$ при всех значениях t . Любую числовую последовательность, в которой аргумент изменяется через равные интервалы, можно также представить в виде решетчатой функции.

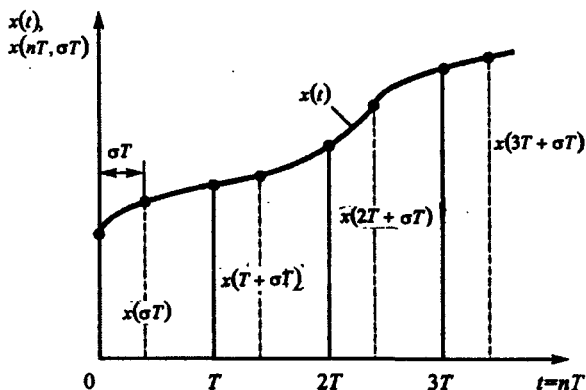


Рис. 6.3. Смещенная решетчатая функция

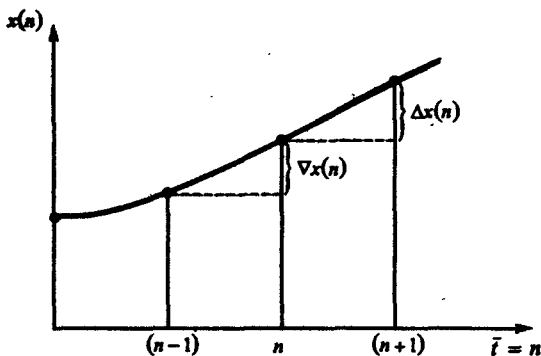


Рис. 6.4. Первая прямая и обратная разности решетчатой функции

Часто оказывается удобным от абсолютного времени t перейти к относительному времени $\bar{t} = t/T$. В этом случае решетчатая функция обозначается так:

$$x(n) = x(\bar{t}) \text{ при } \bar{t} = n,$$

где $n = 0, 1, 2, \dots$, а смещенная решетчатая функция равна

$$x(n, \sigma) = x(\bar{t}) \text{ при } \bar{t} = n + \sigma,$$

где $n = 0, 1, 2, \dots$ и $0 \leq \sigma \leq 1$.

Отметим, что решетчатые функции изменяют свои значения только при целочисленных значениях независимой переменной.

Скорость изменения решетчатой функции характеризуется ее первой разностью. Вообще конечные разности решетчатых функций являются дискретными аналогами производных непрерывных функций. При этом различают прямые и обратные разности (рис. 6.4). Прямая разность первого порядка решетчатой функции $x(n)$ определяется выражением

$$\Delta x(n) = x(n+1) - x(n),$$

а обратная разность первого порядка

$$\nabla x(n) = x(n) - x(n-1).$$

Конечные разности произвольного порядка определяются из рекуррентных выражений

$$\Delta^k x(n) = \Delta [\Delta^{k-1} x(n)] = \Delta^{k-1} x(n+1) - \Delta^{k-1} x(n);$$

$$\nabla^k x(n) = \nabla [\nabla^{k-1} x(n)] = \nabla^{k-1} x(n) - \nabla^{k-1} x(n-1).$$

Для вычисления конечных разностей через значения решетчатой функции $x(n)$ можно использовать формулы:

$$\Delta^k x(n) = \sum_{i=0}^k (-1)^i C_k^i x(n+k-i); \quad (6.1)$$

$$\nabla^k x(n) = \sum_{i=0}^k (-1)^i C_k^i x(n-i), \quad (6.2)$$

где C_k^i — число сочетаний из k элементов по i . Сопоставляя формулы (6.1) и (6.2), получим соотношение между прямой и обратной конечными разностями

$$\nabla^k x(n) = \Delta^k x(n-k).$$

Дискретным аналогом интеграла непрерывной функции в пределах от 0 до t для решетчатой функции является конечная сумма, определяемая по формуле

$$x_{\Sigma}(n) = \sum_{i=0}^{n-1} x(i). \quad (6.3)$$

Для фиксированных n значения x_{Σ} вычисляются элементарно по значениям $x(i)$. Определение же общего выражения для x_{Σ} как функции i в явном виде возможно только для простейших функций $x(n)$. Непосредственно из выражения (6.3) вытекают следующие свойства конечной суммы:

$$\sum_{i=0}^{n-1} x(i) = \sum_{i=n-1}^0 x(i); \quad \sum_{i=0}^{n-1} x(i) = \sum_{i=1}^n x(n-i).$$

Операции вычисления конечных разностей и конечных сумм являются линейными операциями и для них справедлив принцип суперпозиции.

При исследовании непрерывных систем пользуются дифференциальными уравнениями, определяющими связь между непрерывной функцией $x(t)$ и ее производными $d^k x(t)/dt^k$. Аналогично, соотношение между решетчатой функцией $x(n)$ и ее разностями $\Delta^k x(n)$ или $\nabla^k x(n)$ определяет уравнение в конечных разностях или разностное уравнение. Если это соотношение линейно, то разностное уравнение называется *линейным*.

Линейное разностное уравнение с постоянными коэффициентами можно представить в форме

$$a_k \Delta^k x(n) + a_{k-1} \Delta^{k-1} x(n) + \dots + a_0 x(n) = f(n), \quad (6.4)$$

где $f(n)$ — известная решетчатая функция, $x(n)$ — искомая решетчатая функция, представляющая собой решение разностного уравнения. Используя соотношение (6.1), уравнение (6.4) можно записать в виде

$$b_k x(n+k) + b_{k-1} x(n+k-1) + \dots + b_0 x(n) = f(n), \quad (6.5)$$

которое является рекуррентным уравнением, позволяющим вычислить значение $x(n)$ при любом n , если известны текущее значение входа $f(n)$ и k предшествующих значений $x(n-1), \dots, x(n-k)$ выхода. Эта особенность определения решения отличает разностные уравнения от дифференциальных и далее будет использована неоднократно. В том случае, когда $f(n) \equiv 0$, уравнения (6.4) и (6.5) называются *однородными*. Разностное уравнение можно записать не только относительно прямых, но и относительно обратных разностей. Коэффициенты a и b разностных уравнений (6.4), (6.5) связаны соотношением

$$b_l = \sum_{i=0}^l (-1)^{l-i} a_i C_{k-i}^{l-i},$$

где C_{k-i}^{l-i} — число сочетаний из $(k-i)$ элементов по $(l-i)$, $l = 0, 1, 2, \dots, k$. Разностное уравнение в форме (6.5), содержащее $x(n)$ и $x(n+k)$, называют *разностным уравнением k -го порядка*. Порядок разностного уравнения (6.4) в общем случае может не совпадать с порядком наивысшей разности, которая входит в него. Совпадение имеет место только при $b_0 \neq 0$ и $b_k \neq 0$. Классические методы решения разностных уравнений во многом аналогичны классическим методам решения дифференциальных уравнений. Широко используются операторные методы решения разностных уравнений, основанные на переходе от решетчатых функций к их изображениям.

6.2. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ИДЕАЛЬНОГО ИМПУЛЬСНОГО ЭЛЕМЕНТА

Как показано в § 5.3, идеальный импульсный элемент можно рассматривать как импульсный модулятор с несущей в виде последовательности мгновенных единичных импульсов

$$\delta_T(t) = \sum_{n=-\infty}^{+\infty} \delta(t-nT) \quad (6.6)$$

и огибающей в виде входного непрерывного сигнала $x(t)$, т.е.

$$x^*(t) = x(t)\delta_T(t). \quad (6.7)$$

Функцию $\delta_T(t)$ называют *функцией единичных импульсов* и широко используют при исследовании импульсных систем.

Подставляя (6.6) в (6.7) и учитывая, что для физических систем $x(t) \equiv 0$ при $t < 0$, получим

$$x^*(t) = \sum_{n=0}^{\infty} x(nT) \delta(t - nT), \quad (6.8)$$

где $x(nT) = x(t)|_{t=nT}$ — значение входного сигнала $x(t)$ ИИЭ при $t = nT$.

Уравнение (6.8) описывает идеальный импульсный элемент во временной области. Найдем изображение по Лапласу (L -преобразование) функции $x^*(t)$, т.е.

$$X^*(s) = L\{x^*(t)\} = \sum_{n=0}^{\infty} x(nT) L\{\delta(t - nT)\}. \quad (6.9)$$

Учитывая, что

$$L\{\delta(t - nT)\} = \int_0^{\infty} \delta(t - nT) e^{-st} dt = e^{-nsT},$$

из (6.9) получаем

$$X^*(s) = \sum_{n=0}^{\infty} x(nT) e^{-nsT}. \quad (6.10)$$

Это преобразование является дискретным аналогом прямого преобразования Лапласа функции $x(t)$:

$$X(s) = \int_0^{\infty} x(t) e^{-st} dt.$$

Отличие заключается лишь в том, что интеграл в L -преобразовании заменен суммой, а вместо непрерывной функции $x(t)$ в (6.10) фигурирует соответствующая решетчатая функция $x(nT)$.

Таким образом, выражение (6.10) можно рассматривать как преобразование решетчатой функции $x(nT)$ — так называемое прямое дискретное преобразование Лапласа (L_D -преобразование). Нетрудно заметить, что L -преобразование модулированной последовательности $x^*(t)$ мгновенных импульсов на выходе ИИЭ равно дискретному преобразованию Лапласа (L_D -преобразованию) соответствующей решетчатой функции $x(nT)$, т.е.

$$X^*(s) = L\{x^*(t)\} = L_D\{x(nT)\}.$$

Основные свойства дискретного преобразования Лапласа подробно рассмотрены в [29], [88].

Периодическая последовательность δ -импульсов (6.6) может быть представлена в виде комплексного ряда Фурье

$$\delta_T(t) = \sum_{n=-\infty}^{+\infty} c_n e^{jn\omega_0 t},$$

где $\omega_0 = \frac{2\pi}{T}$ — частота квантования по времени; c_n — коэффициенты Фурье:

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} \delta_T(t) e^{-jn\omega_0 t} dt = \frac{1}{T}.$$

Таким образом, получаем второе уравнение, описывающее во временной области преобразование идеальным импульсным элементом непрерывного сигнала $x(t)$:

$$x^*(t) = x(t) \frac{1}{T} \sum_{n=-\infty}^{+\infty} e^{jn\omega_0 t} = \frac{1}{T} \sum_{n=-\infty}^{+\infty} x(t) e^{jn\omega_0 t}.$$

Преобразование Лапласа этого уравнения дает

$$X^*(s) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} L\{x(t) e^{jn\omega_0 t}\} = \frac{1}{T} \sum_{n=-\infty}^{+\infty} X(s + jn\omega_0). \quad (6.11)$$

Уравнение (6.11) устанавливает связь между изображениями Лапласа непрерывной функции $x(t)$ и решетчатой функции $x(nT)$, т.е. между обычным и дискретным преобразованиями Лапласа, и справедливо при выполнении условия $x(t) \equiv 0$ при $t \leq 0$.

Если же $x(0) \neq 0$, то следует пользоваться другим уравнением:

$$X^*(s) = \frac{x(0)}{2} + \frac{1}{T} \sum_{n=-\infty}^{+\infty} X(s + jn\omega_0). \quad (6.12)$$

Операцию нахождения $X^*(s)$ по $X(s)$, определяемую соотношениями (6.11), (6.12) называют D -преобразованием и обозначают

$$X^*(s) = D\{X(s)\}.$$

С помощью D -преобразования легко определить спектральные или частотные характеристики идеального импульсного элемента. Заменяя в (6.11) s на $j\omega$, находим связь между спектрами входного и выходного сигналов идеального импульсного элемента (предполагается, что все полюсы $X^*(s)$ — левые):

$$X^*(j\omega) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} X[j(\omega + n\omega_0)]. \quad (6.13)$$

Уравнение (6.13) означает, что частотный спектр на выходе идеального импульсного элемента представляет собой сумму частотных спек-

тров непрерывного сигнала на входе, смещенных по оси частот на величину $n\omega_0$. Отсюда же следует, что преобразование Лапласа сигнала на выходе ИИЭ является периодической функцией s с периодом, равным ω_0 , т.е.

$$X^*(s) = X^*(s \pm jk\omega_0),$$

где k — целое число. Из того же уравнения (6.13) следует, что частотный спектр сигнала на выходе ИИЭ также является периодической функцией ω с периодом, равным частоте квантования ω_0 , т.е.

$$X^*(j\omega) = X^*[j(\omega \pm k\omega_0)].$$

Периодичность $X^*(j\omega)$ видна из рассмотрения амплитудного спектра $|X^*(j\omega)|$ сигнала на выходе ИИЭ (рис. 6.5), который состоит из спектра $\frac{1}{T}|X(j\omega)|$ непрерывного сигнала на входе и подобных ему дополнительных спектров. Спектр $|X^*(j\omega)|$ полностью определяется диапазоном частот $\left(-\frac{\omega_0}{2}, \frac{\omega_0}{2}\right)$ или, в силу симметрии, — диапазоном $\left(0, \frac{\omega_0}{2}\right)$. Вообще, периодические временные функции имеют дискретные частотные спектры, а дискретные временные функции — периодические частотные спектры.

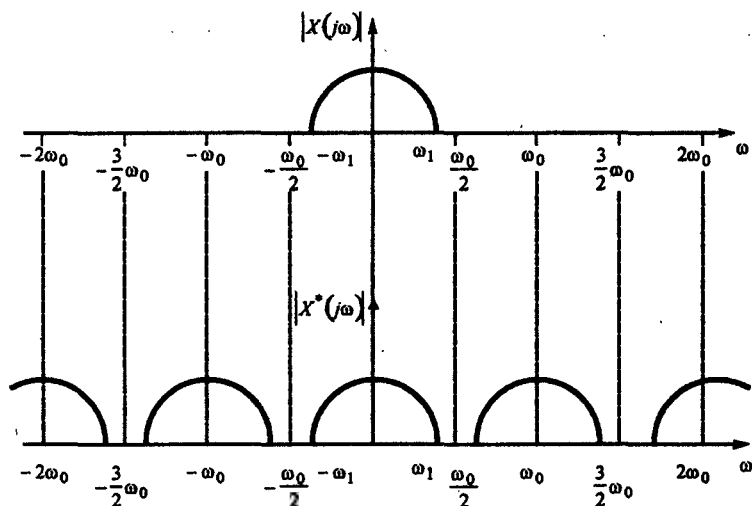


Рис. 6.5. Частотный спектр импульсного сигнала при $\omega_0 > 2\omega_1$

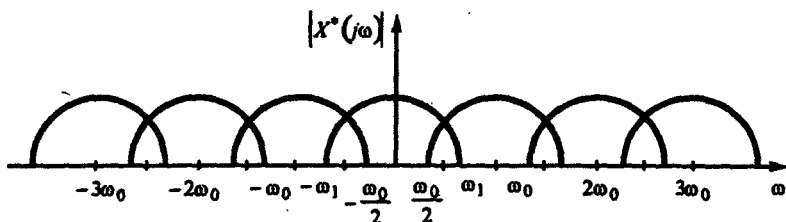


Рис. 6.6. Частотный спектр импульсного сигнала при $\omega_0 < 2\omega_1$

Непрерывный сигнал $x(t)$ может быть теоретически восстановлен по сигналу $x^*(t)$ при условии фильтрации дополнительных (боковых) составляющих спектра $|X^*(j\omega)|$, т.е. с помощью идеального полосового фильтра с амплитудной частотной характеристикой

$$|H(j\omega)| = T, \quad -\omega_1 \leq \omega \leq \omega_1; \quad |H(j\omega)| = 0, \quad \omega < -\omega_1 \text{ и } \omega > \omega_1.$$

Однако, если частота прерывания ω_0 меньше удвоенной наивысшей частоты ω_1 , содержащейся в спектре входного сигнала, основной и дополнительные спектры накладываются так, как показано на рис. 6.6. В этом случае сигнал $x(t)$ не может быть восстановлен фильтрацией сигнала на выходе идеального импульсного элемента.

Таким образом, можно сделать вывод о том, что если непрерывный сигнал $x(t)$ обладает конечным спектром $|X(j\omega)|$ с частотой среза ω_1 , то его квантование по времени с частотой $\omega_0 \geq 2\omega_1$ не приводит к потере информации.

Кроме уравнения (6.11) существует другой способ определения $X^*(s)$ по $X(s)$. На основании теоремы свертки в комплексной области преобразование Лапласа уравнения (6.7) можно записать как свертку функций $X(s)$ и $\Delta_T(s)$, т.е.

$$X^*(s) = X(s) * \Delta_T(s), \quad (6.14)$$

где

$$\Delta_T(s) = \sum_{n=0}^{\infty} e^{-nsT} = 1 + e^{-Ts} + e^{-2Ts} + \dots$$

При $|e^{-sT}| < 1$ этот ряд представляет собой убывающую геометрическую прогрессию и его сумма равна

$$\Delta_T(s) = \frac{1}{1 - e^{-Ts}}. \quad (6.15)$$

С учетом (6.15) уравнение (6.14) можно записать в виде

$$X^*(s) = \frac{1}{2\pi j} \oint_{c_1} X(s-p) \frac{1}{1-e^{-Tp}} dp, \quad (6.16)$$

где c_1 — контур интегрирования на плоскости p , который включает все полюсы $\Delta_r(s)$.

Для вычисления контурного интеграла (6.16) используют теорему о вычетах [88].

В отличие от изображений непрерывных функций, являющихся функциями переменной s , изображения решетчатых функций являются функциями переменной e^{sT} . Введем новую переменную в выражение (6.10)

$$z = e^{sT}$$

и сокращенное обозначение $X(z)$ для получающейся функции от z .

Таким образом,

$$X(z) = \sum_{n=0}^{\infty} x(nT) z^{-n}. \quad (6.17)$$

Выражение (6.17) определяет математическую операцию, получившую название прямого z -преобразования. Функция $X(z)$ называется z -преобразованием от $x(t)$ или, более точно, от $x(nT)$ и символически записывается в виде

$$X(z) = Z\{x(t)\} = Z\{x(nT)\},$$

где Z — оператор z -преобразования. Связь между z -преобразованием и дискретным преобразованием Лапласа можно записать в виде соотношения

$$Z\{x(nT)\} = L\{x^*(t)\} = L_D\{x(nT)\}$$

при $s = \frac{1}{T} \ln z$.

Рассмотрим примеры вычисления z -преобразований некоторых элементарных функций с использованием выражения (6.17).

Пример 1. При исследовании импульсных систем широко используется понятие импульсной решетчатой функции $x(nT) = k\delta_0(nT)$, где $\delta_0(nT)$ — единичная импульсная решетчатая функция, равная

$$\delta_0(nT) = \begin{cases} 1 & \text{при } n = 0; \\ 0 & \text{при } n \neq 0. \end{cases}$$

Функция $\delta_0(nT)$ играет в импульсных системах такую же важную роль, как $\delta(t)$ функция в непрерывных системах.

Импульсной решетчатой функции $x(nT)$ соответствует функция непрерывного времени

$$x(t) = \begin{cases} k & \text{при } t = 0; \\ 0 & \text{при } t \neq 0. \end{cases}$$

Пример 2. Ступенчатая решетчатая функция определяется выражением

$$x(nT) = k1(nT) = k.$$

Используя уравнение (6.17), получим

$$X(z) = \sum_{n=0}^{\infty} kz^{-n} = k(1 + z^{-1} + z^{-2} + \dots) = \frac{kz}{z-1}.$$

Пример 3. Линейная функция непрерывного времени описывается так:

$$x(t) = kt.$$

Из этого выражения при $t = nT$ имеем

$$x(nT) = knT. \quad (6.18)$$

Подставляя (6.18) в (6.17), получим

$$X(z) = \sum_{n=0}^{\infty} knTz^{-n} = kT(z^{-1} + 2z^{-2} + 3z^{-3} + \dots).$$

При $|z^{-1}| < 1$ этот ряд сходится и можно записать изображение линейной функции

$$X(z) = \frac{kTz}{(z-1)^2}.$$

Пример 4. Экспоненциальная функция непрерывного времени $x(t) = ke^{-\alpha t}$ при $t = nT$ запишется так:

$$x(nT) = ke^{-\alpha nT}.$$

Подставляя это выражение в (6.17), получим

$$X(z) = \sum_{n=0}^{\infty} ke^{-\alpha nT} z^{-n} = k(1 + e^{-\alpha T} z^{-1} + e^{-2\alpha T} z^{-2} + \dots).$$

При $|z^{-1}| < 1$ это выражение представляет убывающую геометрическую прогрессию, сумма которой равна

$$X(z) = \frac{kz}{z - e^{-\alpha T}}.$$

Для нахождения z -изображения по заданному оригиналу $x(t)$ или $x(nT)$ можно использовать специальные таблицы соответствия между функциями времени и их z -изображениями. Такие таблицы приведены в [29]. Изображения некоторых функций даны в табл. 6.1.

$x(t)$	$X(s)$	$X(z)$	$X(z, \sigma)$
$x(t) = \begin{cases} 1 & \text{при } t = 0 \\ 0 & \text{при } t \neq 0 \end{cases}$	—	1	—
$x(t) = \begin{cases} 1 & \text{при } t = kT \\ 0 & \text{при } t \neq kT \end{cases}$	—	z^{-k}	—
$1(t)$	$\frac{1}{s}$	$\frac{z}{z-1}$	$\frac{z}{z-1}$
t	$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$	$\frac{Tz}{(z-1)^2} + \frac{T\sigma z}{z-1}$
$\frac{1}{2}t^2$	$\frac{1}{s^3}$	$\frac{T^2 z(z+1)}{2(z-1)^3}$	$\frac{T^2 z}{2} \left[\frac{\sigma^2}{z-1} + \frac{2\sigma+1}{(z-1)^2} + \frac{2}{(z-1)^3} \right]$
$e^{-\alpha t}$	$\frac{1}{s+\alpha}$	$\frac{z}{z-d}, d = e^{-\alpha T}$	$\frac{z}{z-d} d^\sigma, d = e^{-\alpha T}$
$1 - e^{-\alpha t}$	$\frac{\alpha}{s(s+\alpha)}$	$\frac{z(1-d)}{(z-1)(z-d)},$ $d = e^{-\alpha T}$	$\frac{z}{z-1} - \frac{z d^\sigma}{z-d},$ $d = e^{-\alpha T}$
$e^{-at} - e^{-bt}$	$\frac{b-a}{(s+a)(s+b)}$	$\frac{z}{z-d_1} - \frac{z}{z-d_2},$ $d_1 = e^{-aT}, d_2 = e^{-bT}$	$\frac{z d_1^\sigma}{z-d_1} - \frac{z d_2^\sigma}{z-d_2}$
$\sin \omega t$	$\frac{\omega}{s^2 + \omega^2}$	$\frac{z \sin \omega T}{z^2 - 2z \cos \omega T + 1}$	$\frac{z^2 \sin \sigma \omega T + z \sin(1-\sigma)\omega T}{z^2 - 2z \cos \omega T + 1}$
$\cos \omega t$	$\frac{s}{s^2 + \omega^2}$	$\frac{z^2 - z \cos \omega T}{z^2 - 2z \cos \omega T + 1}$	$\frac{z^2 \cos \sigma \omega T - z \cos(1-\sigma)\omega T}{z^2 - 2z \cos \omega T + 1}$

Основные теоремы и свойства z -преобразования аналогичны свойствам L -преобразования и рассмотрены в [29], [77]. Остановимся лишь на некоторых из них.

- *Линейность.* Если функции $x_1(nT)$, $x_2(nT)$ и $x(nT)$ преобразуемы по z и имеют z -преобразования, соответственно равные $X_1(z)$, $X_2(z)$ и $X(z)$, и a — постоянный коэффициент, то

$$Z\{ax(nT)\} = aX(z), \quad Z\{x_1(nT) \pm x_2(nT)\} = X_1(z) \pm X_2(z).$$

- *Сдвиг во временной области.* Если функция $x(nT)$ преобразуема по z и имеет z -преобразование $X(z)$, то при сдвиге решетчатой функции вправо (запаздывание) имеем

$$Z\{x(nT - kT)\} = z^{-k} X(z),$$

где k — неотрицательное целое число; $x(nT) = 0$ для $n < 0$. При сдвиге влево (упреждение) получим

$$Z\{x(nT + kT)\} = z^k \left[X(z) - \sum_{i=0}^{k-1} x(iT) z^{-i} \right].$$

Если $x(nT) = 0$ при $n = 0, 1, \dots, k-1$, то последнее выражение упрощается:

$$Z\{x(nT + kT)\} = z^k X(z).$$

- *Конечное значение.* Если функци $x(nT)$ соответствует z -преобразование $X(z)$, которое не имеет полюсов $|z| \geq 1$, то

$$\lim_{z \rightarrow 1} \frac{z-1}{z} X(z) = \lim_{n \rightarrow \infty} x(nT).$$

- *Начальное значение.* Если функция $x(nT)$ имеет z -преобразование $X(z)$ и предел $\lim_{z \rightarrow \infty} X(z)$ существует, то

$$\lim_{z \rightarrow \infty} X(z) = \lim_{n \rightarrow 0} x(nT).$$

- *Изменение масштаба в области z .* Если функция $x(nT)$ преобразуема по z и имеет z -преобразование $X(z)$, то

$$Z\{e^{\pm \alpha nT} x(nT)\} = X(z e^{\mp \alpha T}),$$

где α — целое число.

6.3. УРАВНЕНИЯ И ИМПУЛЬСНАЯ ПЕРЕДАТОЧНАЯ ФУНКЦИЯ РАЗОМКНУТОЙ ИМПУЛЬСНОЙ СИСТЕМЫ

Выход приведенной непрерывной части импульсной системы (см. рис. 5.13) представляет собой непрерывный сигнал, описываемый функцией времени $y(t)$. Для упрощения анализа принято рассматривать этот сигнал в дискретные моменты времени, совпадающие с моментами замыкания идеального импульсного элемента на входе. Это равносильно включению фиктивного идеального импульсного элемента на выходе системы, работающего синхронно и синфазно с основным импульсным элементом и рассмотрению в качестве выходной переменной функции $y^*(t)$.

Используя выражение (6.10), определим L -преобразование выхода импульсной системы

$$Y^*(s) = L\{y^*(t)\} = \sum_{n=0}^{\infty} y(nT) e^{-nsT}. \quad (6.19)$$

Значения $y(nT)$ сигнала на выходе ПНЧ для $t = nT$ определяются из выражения для функции $y(t)$, имеющего вид:

$$y(t) = \sum_{i=0}^{\infty} w(t - iT) x(iT),$$

где $w(t)$ — импульсная переходная или весовая функция приведенной непрерывной части с передаточной функцией $W(s)$.

Следовательно, значения выхода в моменты времени $t = nT$ равны

$$y(nT) = \sum_{i=0}^{\infty} w(nT - iT) x(iT). \quad (6.20)$$

Подставляя (6.20) в (6.19), получим

$$Y^*(s) = \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} w(nT - iT) x(iT) e^{-nsT}. \quad (6.21)$$

Подстановкой $m = n - i$ и $n = i + m$ уравнение (6.21) приводится к виду

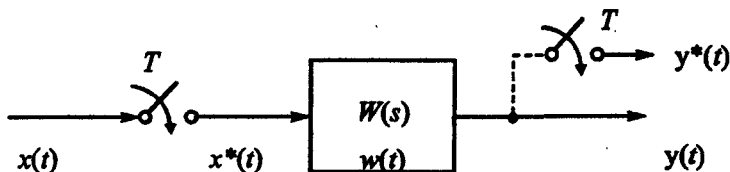


Рис. 6.7. К определению импульсной передаточной функции разомкнутой импульсной системы

$$Y^*(s) = \sum_{i=0}^{\infty} x(iT) \sum_{m=-i}^{\infty} w(mT) e^{-isT} e^{-msT}.$$

Учитывая, что

$$w(mT) \equiv 0 \text{ для } m < 0,$$

окончательно получим

$$Y^*(s) = \sum_{i=0}^{\infty} x(iT) e^{-isT} \sum_{m=0}^{\infty} w(mT) e^{-msT}. \quad (6.22)$$

Исходя из определения L_D -преобразования (6.10), можно привести уравнение (6.22) к виду

$$Y^*(s) = X^*(s)W^*(s), \quad (6.23)$$

где

$$W^*(s) = \frac{Y^*(s)}{X^*(s)} = \sum_{m=0}^{\infty} w(mT) e^{-msT} = L_D \{w(mT)\} \quad (6.24)$$

— импульсная передаточная функция разомкнутой системы в s -форме (так называемая импульсная передаточная функция со звездочкой). Таким образом, импульсная передаточная функция в s -форме является отношением дискретных преобразований Лапласа выхода и входа при нулевых начальных условиях.

Путем подстановки $z = e^{sT}$ в (6.23) и (6.24) можно получить уравнения для z -изображений, т.е.

$$Y(z) = X(z)W(z)$$

и

$$W(z) = \frac{Y(z)}{X(z)} = \sum_{m=0}^{\infty} w(mT) z^{-m}. \quad (6.25)$$

Здесь $W(z)$ — импульсная передаточная функция в z -форме. Следовательно, импульсная передаточная функция в z -форме может быть определена как отношение z -изображения импульсного выхода системы к изображению импульсного входа при нулевых начальных условиях. Выражение (6.25) показывает, что импульсная передаточная функция представляет z -преобразование импульсной переходной функции приведенной непрерывной части системы, т.е.

$$W(z) = Z\{w(t)\} = Z\{w(nT)\} \quad (6.26)$$

или

$$W(z) = Z\{W(s)\},$$

в предположении, что z -преобразованию подвергается импульсная переходная функция $w(t) = L^{-1}\{W(s)\}$ или соответствующая ей решетчатая функция $w(nT)$.

Импульсную передаточную функцию $W^*(s)$ можно также определить с помощью D -преобразования через передаточную функцию приведенной непрерывной части системы $W(s)$ на основании выражений (6.11) и (6.12) в виде

$$W^*(s) = \frac{1}{T} \sum_{n=-\infty}^{+\infty} W(s + jn\omega_0) \quad (6.27)$$

при $w(t) \equiv 0, t \leq 0$;

$$W^*(s) = \frac{w(0)}{2} + \frac{1}{T} \sum_{n=-\infty}^{+\infty} W(s + jn\omega_0) \quad (6.28)$$

при $w(0) \neq 0$.

Таким образом, импульсная передаточная функция представляет собой D -преобразование передаточной функции ее приведенной непрерывной части

$$W^*(s) = D\{W(s)\}.$$

Во многих случаях такой способ вычисления импульсной передаточной функции оказывается наиболее простым. Можно также использовать и интегральное преобразование (6.16). В этом случае оно приобретает вид

$$W^*(s) = \frac{1}{2\pi j} \oint_{\sigma_1} W(s-p) \frac{1}{1-e^{-Tp}} dp. \quad (6.29)$$

Контурный интеграл находится по теореме о вычетах [88].

Рассмотрим примеры определения импульсных передаточных функций системы (рис. 6.8) с формирующим элементом типа фиксатора (см. § 5.3), передаточная функция которого $H(s)$ имеет вид (5.1). В соответствии с (6.26) импульсная передаточная функция системы определяется выражением

$$W(z) = Z\{H(s)W(s)\} = Z\left\{\frac{1-e^{-Ts}}{s}W(s)\right\}.$$

Сделав замену $e^{sT} = z$ и обозначив:

$$\frac{W(s)}{s} = W_1(s); \quad Z\{W_1(s)\} = W_1(z),$$

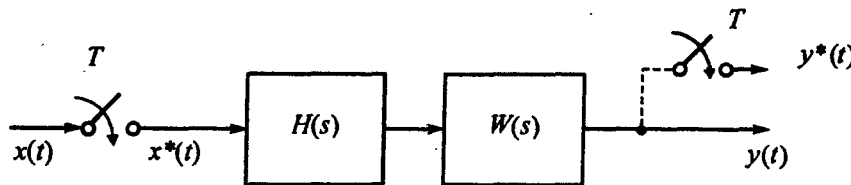


Рис. 6.8. Структурная схема амплитудно-импульсной системы с модуляцией первого рода

Окончательно имеем

$$W(z) = (1 - z^{-1})Z\{W_1(s)\} = \frac{z-1}{z}W_1(z), \quad (6.30)$$

где

$$W_1(z) = \sum_{n=0}^{\infty} w_1(nT)z^{-n}; \quad (6.31)$$

$$w_1(nT) = w_1(t)|_{t=nT} = L^{-1}\{W_1(s)\}_{t=nT}. \quad (6.32)$$

Пример 5. Непрерывная часть представляет собой интегрирующее звено с передаточной функцией

$$W(s) = \frac{k}{s}.$$

В этом случае, $w_1(s) = \frac{k}{s^2}$ а $w_1(nT) = L^{-1}\left\{\frac{k}{s^2}\right\} = kt|_{t=nT} = knT$.

Тогда $w_1(z) = \sum_{n=0}^{\infty} knTz^{-n} = \frac{kTz}{(z-1)^2}$ и, окончательно, имеем из формулы (6.30) выражение для импульсной передаточной функции

$$W(z) = \frac{kT}{z-1}. \quad (6.33)$$

Пример 6. Непрерывная часть является аperiодическим звеном с передаточной функцией

$$W(s) = \frac{k}{1 + sT_1}.$$

В этом случае

$$W_1(s) = \frac{k}{s(1 + sT_1)} = k\left(\frac{1}{s} - \frac{1}{s + \alpha_1}\right), \quad \alpha_1 = \frac{1}{T_1}.$$

Из выражений (6.32), (6.31) определяем

$$\begin{aligned} w_1(nT) &= k \left[1(nT) - e^{-\alpha_1 nT} \right]; \\ W_1(z) &= k \left(\frac{z}{z-1} - \frac{z}{z-d} \right), \quad d = e^{-\alpha_1 T}. \end{aligned} \quad (6.34)$$

Подставляя (6.34) в (6.30), получаем импульсную передаточную функцию системы

$$W(z) = \frac{k(1-d)}{z-d}. \quad (6.35)$$

Пример 7. Непрерывная часть представляет собой фильтр с передаточной функцией

$$W(s) = \frac{k}{s(1+sT_1)(1+sT_2)\dots(1+sT_n)}. \quad (6.36)$$

Представим (6.36) в виде

$$W(s) = \frac{k}{s} + \sum_{i=1}^n \frac{C_i k}{1+T_i s}, \quad i = 1, 2, \dots, n.$$

Используя результаты, полученные в примерах 5, 6, записываем выражение для импульсной передаточной функции:

$$W(z) = \frac{kT}{z-1} + k \sum_{i=1}^n \frac{C_i(1-d_i)}{z-d_i}, \quad d_i = e^{-T/T_i}. \quad (6.37)$$

Пример 8. Передаточная функция непрерывной части разомкнутой импульсной системы имеет вид

$$W(s) = \frac{k(1+\tau_1 s)(1+\tau_2 s) \dots (1+\tau_m s)}{(1+T_1 s)(1+T_2 s) \dots (1+T_n s)}.$$

Разложим ее на простые дроби:

$$W(s) = k \sum_{i=1}^n \frac{C_i}{1+T_i s} = k \sum_{i=1}^n \frac{C'_i}{s + \alpha_i},$$

где $\alpha_i = 1/T_i$, а C'_i и $C_i = C'_i T_i$ — коэффициенты, определяемые в соответствии с теоремой разложения.

Импульсная передаточная функция системы определяется выражением

$$W(z) = \frac{z-1}{z} Z \left\{ \frac{W(s)}{s} \right\} = \frac{z-1}{z} Z \left\{ k \sum_{i=1}^n \frac{C_i}{s(1+T_i s)} \right\}.$$

Используя результат, полученный в примере 6, записываем:

$$W(z) = k \sum_{i=1}^n \frac{C'_i}{\alpha_i} \frac{1-d_i}{z-d_i} = k \sum_{i=1}^n \frac{C_i(1-d_i)}{z-d_i}, \quad d_i = e^{-\alpha_i T} = e^{-T/T_i}.$$

Пример 9. Непрерывная часть астатической системы имеет передаточную функцию:

$$W(s) = \frac{k \prod_{j=1}^m (1 + \tau_j s)}{s^v \prod_{i=1}^n (1 + T_i s)}.$$

Разложение на простые дроби для $v = 1$ имеет вид:

$$W(s) = \frac{k}{s} + k \sum_{i=1}^n \frac{C_i}{1 + T_i s},$$

а соответствующая импульсная передаточная функция системы с астатизмом первого порядка (получена по результатам примеров 5, 8) определяется выражением

$$W(z) = \frac{kT}{z-1} + k \sum_{i=1}^n \frac{C_i(1-d_i)}{z-d_i}, \quad d_i = e^{-T/T_i}. \quad (6.38)$$

Для непрерывной части с двумя интегрирующими звеньями ($v = 2$) в передаточной функции $W(s)$ — системы с астатизмом второго порядка разложение на простые дроби имеет вид

$$W(s) = \frac{k}{s^2} + \frac{k_Q}{s} + k \sum_{i=1}^n \frac{C_i}{1 + T_i s},$$

где $k_Q = k \left(\sum_{j=1}^m \tau_j - \sum_{i=1}^n T_i \right)$ — так называемая условная добротность по скорости.

Импульсная передаточная функция такой системы определяется выражением

$$W(z) = \frac{kT^2}{2} \frac{z+1}{(z-1)^2} + \frac{k_Q T}{z-1} + k \sum_{i=1}^n \frac{C_i(1-d_i)}{z-d_i}. \quad (6.39)$$

Все рассмотренные примеры 5...9 соответствовали импульсной системе, приведенной на рис. 6.8 с формирующим элементом типа экстраполятора нулевого порядка.

Для того чтобы определить импульсную передаточную функцию такой системы с формирующим элементом произвольного типа, необходимо:

1. Определить передаточную функцию ПНЧ: $W_{\text{пнч}}(s) = H(s)W(s)$.
2. С помощью обратного преобразования Лапласа найти импульсную переходную функцию ПНЧ: $w(t) = L^{-1}\{W_{\text{пнч}}(s)\}$.
3. Определить весовую последовательность системы (решетчатую функцию веса) $w(nT) = w(t)|_{t=nT}$.

4. Найти сумму ряда в правой части выражения $W(z) = \sum_{n=0}^{\infty} w(nT)z^{-n}$.

Если передаточная функция приведенной непрерывной части системы $W_{\text{пч}}(s)$ является сложной, то ее необходимо представить в виде суммы простых дробей и для каждой из них последовательно выполнить этапы 2, 3 и 4. Искомая импульсная передаточная функция равна

$$W(z) = \sum_k W_k(z).$$

Применимость рассмотренного подхода зависит от возможности вычисления суммы бесконечного ряда на этапе 4, который для большинства реальных систем сходится. В том случае, когда выражение для этой суммы в замкнутой форме получить невозможно, импульсная передаточная функция может быть определена по формулам (6.27), (6.28) или (6.29).

В общем виде импульсную передаточную функцию можно записать так:

$$W(z) = \frac{B(z)}{A(z)} = \frac{k}{(z-1)^v} \frac{B_1(z)}{A_1(z)},$$

где $A(z)$, $B(z)$ — полиномы степеней n и m , причем $m \leq n$; k — коэффициент передачи системы; $A_1(z)$, $B_1(z)$ — нормированные полиномы степеней $n-v$ и m , такие, что $B_1(0)/A_1(0) = 1$.

Импульсную передаточную функцию можно характеризовать следующими числовыми показателями:

- порядком системы n , который определяется числом полюсов или степенью знаменателя импульсной передаточной функции $W(z)$;
- порядком астатизма v — числом полюсов $W(z)$, равных единице;
- числом нулей m или степенью полинома числителя импульсной передаточной функции.

Рассмотрим теперь некоторые важные свойства импульсных передаточных функций, которые могут быть полезны при анализе и синтезе импульсных систем.

1. Импульсная передаточная функция в s -форме $W^*(s) = W(z)|_{z=e^{sT}}$ является периодической функцией с периодом $j\omega_0$; $\omega_0 = \frac{2\pi}{T}$ — частота квантования. Свойство периодичности следует из уравнений (6.27), (6.28).

2. Значения импульсной передаточной функции $W(z)$ всегда действительны при $z = 1 (\omega = 0)$ и $z = -1 (\omega = \frac{n\omega_0}{2})$; конечны при $z = 1$, если $W(s)$ (см. выражения (6.27), (6.28)) не имеет полюса в начале координат.

3. Полюсы z , импульсной передаточной функции $W(z)$ связаны с полюсами s , передаточной функции $W(s)$ приведенной непрерывной части системы соотношением $z_i = e^{s_i T}$. Число полюсов $W(z)$ равно числу полюсов $W(s)$, следовательно, степень знаменателя передаточной функции $W(z)$ равна степени знаменателя $W(s)$.

6.4. ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ ИМПУЛЬСНЫХ СИСТЕМ

Физический смысл частотных характеристик импульсных и непрерывных систем очень близок. Особенностью этих характеристик для импульсных систем является то, что они устанавливают связь между гармоническими последовательностями (гармоническими решетчатыми функциями) на входе и выходе импульсного фильтра с передаточной функцией $W^*(s)$ или $W(z)$. Огибающие решетчатых функций изменяются по гармоническому закону.

Если на вход линейного импульсного фильтра подается гармоническая последовательность $x(nT) = A_x \sin \omega nT$, то после окончания переходного процесса на выходе будем иметь также гармоническую последовательность $y(nT) = A_y \sin(\omega nT + \varphi)$.

Заметим, что в отличие от непрерывной гармонической функции гармоническая решетчатая функция в общем случае не является периодической функцией n . Кроме того, амплитуды A_x, A_y не обязательно являются теми максимальными значениями, которых могут достигать те и иные члены соответствующих последовательностей $x(nT), y(nT)$. Амплитуды всегда будут определять лишь верхние границы, но не обязательно максимумы членов этих последовательностей.

Если исходная информация о системе представлена импульсной передаточной функцией $W^*(s)$ или $W(z)$, то для перехода к частотным характеристикам используются замены аргументов $s = j\omega$ или $z = e^{j\omega T}$.

В результате такой замены аргумента получаем комплексный коэффициент передачи импульсной системы

$$W^*(j\omega) = W(e^{j\omega T}). \quad (6.40)$$

Пусть импульсная передаточная функция имеет вид

$$W(z) = \frac{b_m z^m + b_{m-1} z^{m-1} + \dots + b_0}{a_n z^n + a_{n-1} z^{n-1} + \dots + a_0} = \frac{B(z)}{A(z)}.$$

Сделав замену $z = e^{j\omega T}$, получим комплексный коэффициент передачи

$$W(e^{j\omega T}) = \frac{b_m e^{jm\omega T} + b_{m-1} e^{j(m-1)\omega T} + \dots + b_0}{a_n e^{jn\omega T} + a_{n-1} e^{j(n-1)\omega T} + \dots + a_0}. \quad (6.41)$$

Комплексный коэффициент передачи импульсной системы может быть представлен так:

$$W^*(j\omega) = P^*(\omega) + jQ^*(\omega) = R^*(\omega) e^{j\varphi^*(\omega)},$$

где $P^*(\omega)$, $Q^*(\omega)$, $R^*(\omega)$, $\varphi^*(\omega)$ — соответственно вещественная, мнимая, амплитудная и фазовая частотные характеристики импульсной системы. Очевидно

$$R^*(\omega) = \sqrt{P^{*2}(\omega) + Q^{*2}(\omega)}, \quad \varphi^*(\omega) = \text{Arctg} \frac{Q^*(\omega)}{P^*(\omega)} + k\pi, \quad k = 0, \pm 1, \pm 2, \dots;$$

$$P^*(\omega) = R^*(\omega) \cos \varphi^*(\omega), \quad Q^*(\omega) = R^*(\omega) \sin \varphi^*(\omega).$$

При фиксированном значении ω комплексный коэффициент передачи (6.40) изображается вектором на плоскости (P^*, jQ^*) . При изменении ω конец вектора $W^*(j\omega)$ прочерчивает некоторую кривую, которую называют *годографом комплексного коэффициента передачи системы* или ее *амплитудно-фазовой характеристикой*.

Отметим основные особенности частотных характеристик импульсных систем, которые вытекают из свойств импульсной передаточной функции.

1. Частотные характеристики импульсных систем являются периодическими функциями относительно частоты ω с периодом повторения

$\omega_0 = \frac{2\pi}{T}$. Это означает, что при построении этих характеристик достаточно ограничиться изменением ω в диапазоне шириной $\frac{2\pi}{T}$, напри-

мер от $-\frac{\pi}{T}$ до $\frac{\pi}{T}$. Если же учесть, что участки частотной характеристики

в диапазонах ω от $-\frac{\pi}{T}$ до 0 и от 0 до $\frac{\pi}{T}$ симметричны (поскольку

$W^*(j\omega)$ и $W^*(-j\omega)$ — комплексные сопряженные функции), то можно ограничиться построением частотной характеристики в интервале изменения ω от 0 до $\frac{\pi}{T}$.

2. Амплитудно-фазовые частотные характеристики импульсной системы заканчиваются на вещественной оси, так как для $\omega = \frac{\pi}{T}$ комплексный коэффициент передачи (6.41) всегда является действительным числом.

Свойство периодичности частотных характеристик импульсных систем физически объясняется стробоскопическим эффектом, который проявляется в том, что гармоническая решетчатая функция $x(nT) = A_x \sin \omega_0 nT$ на входе импульсного фильтра не изменяется при изменении частоты ω огибающей на любую величину, кратную ω_0 , т.е. последовательность $x(nT)$ будет одной и той же при всех частотах огибающей, равных $\omega + k\omega_0$, $k = 0, 1, 2, \dots$ (рис. 6.9).

Амплитудно-фазовую частотную характеристику импульсной системы можно построить как по выражению для импульсной передаточной функции $W^*(s)$ или $W(z)$, так и по передаточной функции приведенной непрерывной части (с использованием выражений (6.27), (6.28) при $s = j\omega$) или ее импульсной переходной функции (на основе выражения (6.24) при $s = j\omega$).

Рассмотрим примеры построения амплитудно-фазовых частотных характеристик импульсных систем (см. рис. 6.8) с формирующим элементом $H(s)$ типа экстраполятора нулевого порядка.

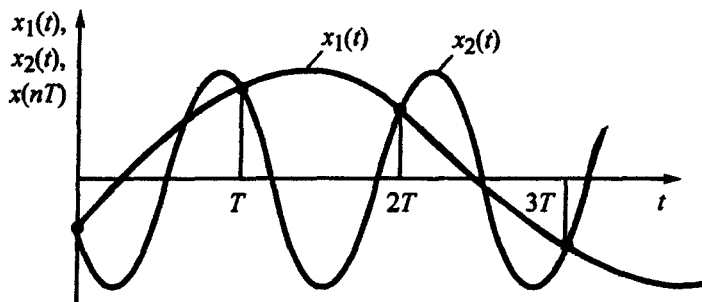


Рис. 6.9. Гармонические сигналы различных частот, соответствующие одной и той же решетчатой функции

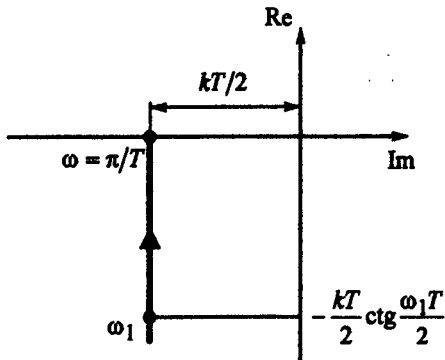


Рис. 6.10. Амплитудно-фазовая частотная характеристика импульсной системы с интегратором

Пример 10. Непрерывная часть представляет собой интегрирующее звено.

Импульсная передаточная функция системы имеет вид $W(z) = \frac{kT}{z-1}$, а комплекс-

ный коэффициент передачи $W^*(j\omega) = \frac{kT}{e^{j\omega T} - 1}$.

Используя формулу Эйлера $e^{j\omega T} = \cos \omega T + j \sin \omega T$, преобразуем последнее выражение к виду

$$W^*(j\omega) = \frac{kT}{\cos \omega T - 1 + j \sin \omega T} = -\frac{kT}{2} - j \frac{kT}{2} \operatorname{ctg} \frac{\omega T}{2}.$$

Как видно из рис. 6.10, амплитудно-фазовая частотная характеристика импульсной системы представляет из себя прямую, параллельную мнимой оси и расположенную в третьем квадранте.

Пример 11. Непрерывная часть системы представляет собой аperiodическое звено с постоянной времени T_1 .

Импульсная передаточная функция определяется выражением

$$W(z) = \frac{k(1-d)}{z-d}, \quad d = e^{-\frac{T}{T_1}}.$$

Комплексный коэффициент передачи в этом случае равен

$$W^*(j\omega) = \frac{k(1-d)}{e^{j\omega T} - d}.$$

Представим последнее выражение в виде

$$W^*(j\omega) = \frac{k(1-d)}{\sqrt{1-2d \cos \omega T + d^2}} e^{-j \arctg \frac{\sin \omega T}{\cos \omega T - d}}.$$

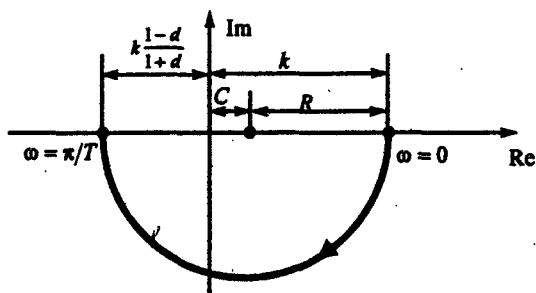


Рис. 6.11. Амплитудно-фазовая частотная характеристика импульсной системы с аperiodическим звеном

На рис. 6.11 представлена амплитудно-фазовая частотная характеристика, соответствующая этому выражению. Как видно из рисунка, она представляет из себя полуокружность, центр которой лежит на вещественной оси на расстоянии от начала координат $C = \frac{kd}{1+d}$ с радиусом, равным $R = \frac{k}{1+d}$.

6.5. ЛОГАРИФМИЧЕСКИЕ ЧАСТОТНЫЕ ХАРАКТЕРИСТИКИ ИМПУЛЬСНЫХ СИСТЕМ

Применение метода логарифмических частотных характеристик для расчета импульсных систем основано на так называемом билинейном преобразовании

$$z = \frac{1+w}{1-w} \quad (6.42)$$

или, соответственно,

$$w = \frac{z-1}{z+1}, \quad (6.43)$$

т.е. на переходе от z -преобразования к w -преобразованию.

Сделав подстановку $z = e^{j\omega T}$, получаем из (6.43)

$$w = \frac{e^{j\omega T} - 1}{e^{j\omega T} + 1} = j \operatorname{tg} \frac{\omega T}{2} = j\bar{\lambda},$$

где $\bar{\lambda} = \operatorname{tg} \frac{\omega T}{2}$ — так называемая относительная (безразмерная) псевдочастота. Обычно вводится в рассмотрение абсолютная псевдочастота

$$\lambda = \frac{2}{T} \bar{\lambda} = \frac{2}{T} \operatorname{tg} \frac{\omega T}{2}.$$

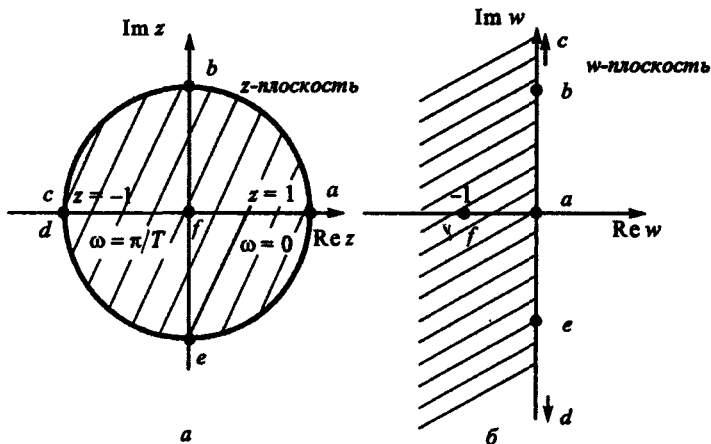


Рис. 6.12. Отображение круга единичного радиуса плоскости z на плоскость w

Функция (6.42) однозначно отображает внутреннюю область единичного круга (рис. 6.12, а) плоскости z в левую полуплоскость (рис. 6.12, б) плоскости w , при этом окружности единичного радиуса на плоскости z соответствует мнимая ось на плоскости w . Таким образом, плоскость w для импульсных систем является аналогом плоскости s для непрерывных систем.

Известно, что основное достоинство метода логарифмических частотных характеристик для непрерывных систем состоит в простоте построения самих характеристик и большой наглядности при их использовании для расчета. Переход к w -преобразованию позволяет распространить этот метод на импульсные системы.

Для перехода от импульсной передаточной функции $W(z)$ к частотной характеристике $W(j\lambda)$ следует сделать подстановку:

$$z = \frac{1+w}{1-w} \bigg|_{w=j\lambda} = \frac{1+j\lambda \frac{T}{2}}{1-j\lambda \frac{T}{2}}. \quad (6.44)$$

Рассмотрим примеры построения логарифмических частотных характеристик импульсной системы (см. рис. 6.8) с формирующим элементом $H(s)$ типа фиксатора.

Комплексный коэффициент передачи $W(j\lambda)$ системы с интегрирующим звеном, импульсная передаточная функция которой имеет вид (6.33), определяется выражением

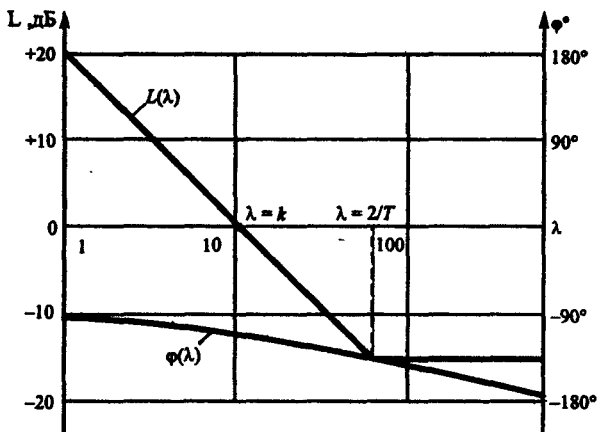


Рис. 6.13. Логарифмические частотные характеристики импульсной системы с интегратором

$$W(j\lambda) = \frac{kT}{\left[\left(1 + j\lambda \frac{T}{2} \right) / \left(1 - j\lambda \frac{T}{2} \right) \right] - 1} = \frac{k \left(1 - j\lambda \frac{T}{2} \right)}{j\lambda}$$

На рис. 6.13 показаны соответствующие логарифмические частотные характеристики

$$L(\lambda) = 20 \lg \left| \frac{k}{j\lambda} \right| + 20 \lg \left| 1 - j\lambda \frac{T}{2} \right|, \quad \varphi(\lambda) = -\frac{\pi}{2} - \operatorname{arctg} \lambda \frac{T}{2}.$$

Как видно из рис. 6.13, в области высоких частот частотные характеристики импульсной системы значительно отличаются от характеристик ее непрерывной части. В области $\lambda > \frac{2}{T}$ фильтрующие свойства интегрирующего звена практически не проявляются.

В области низких частот частотные характеристики импульсной системы достаточно близки частотным характеристикам непрерывной части. Для этой области, или, точнее, при $\omega < \frac{2}{T}$ имеем $\operatorname{tg} \frac{\omega T}{2} \approx \frac{\omega T}{2}$, $\lambda = \frac{2}{T} \operatorname{tg} \frac{\omega T}{2} \approx \omega$, т.е. численные значения частоты ω и абсолютной псевдо-частоты λ практически совпадают.

Если непрерывная часть представляет собой апериодическое звено, то из (6.35) после подстановки (6.44) получим

$$W(j\lambda) = \frac{k(1-d)}{\left[\left(1+j\lambda\frac{T}{2}\right) / \left(1-j\lambda\frac{T}{2}\right)\right] - d} = \frac{k(1-d)\left(1-j\lambda\frac{T}{2}\right)}{1-d+(1+d)j\lambda\frac{T}{2}}, \quad d = e^{-\frac{T}{T_1}}.$$

После несложных преобразований имеем

$$W(j\lambda) = \frac{k\left(1-j\lambda\frac{T}{2}\right)}{1+j\lambda\tau},$$

где принято $\tau = \frac{T}{2} \operatorname{cth} \frac{T}{2T_1}$. Если $T < 2T_1$, то $\operatorname{cth} \frac{T}{2T_1} \approx \frac{2T_1}{T}$ и, следовательно, $\tau \approx T_1$.

Комплексный коэффициент передачи в этом случае равен

$$W(j\lambda) = \frac{k\left(1-j\lambda\frac{T}{2}\right)}{1+j\lambda T_1}. \quad (6.45)$$

Если $T > 2T_1$, то $\operatorname{cth} \frac{T}{2T_1} \approx 1$ и $\tau \approx \frac{T}{2}$. Комплексный коэффициент передачи определяется выражением

$$W(j\lambda) = \frac{k\left(1-j\lambda\frac{T}{2}\right)}{1+j\lambda\frac{T}{2}}. \quad (6.46)$$

На рис. 6.14 показаны логарифмические частотные характеристики, соответствующие выражению (6.45). Как видно из рисунка, в области частот $\lambda < \frac{2}{T}$ частотные характеристики импульсной системы совпадают с частотными характеристиками ее непрерывной части. В области высоких частот, как и в системе с интегрирующим звеном, фильтрующие свойства апериодического звена практически не проявляются. Нетрудно заметить, что при относительно больших значениях периода квантования $T > 2T_1$, когда комплексный коэффициент передачи импульсной системы определяется выражением (6.46), ее частотные характеристики приближаются к частотным характеристикам непрерывной части только при $\lambda \approx \omega \rightarrow 0$.

Построим теперь частотные характеристики импульсной системы, передаточная функция непрерывной части которой определяется выра-

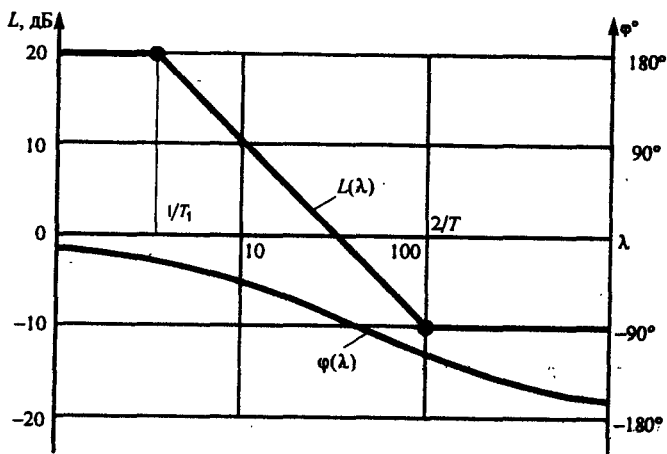


Рис. 6.14. Логарифмические частотные характеристики импульсной системы с аperiodическим звеном

жением (6.36). Комплексный коэффициент передачи системы на основании (6.37) равен

$$W(j\lambda) = \frac{k \left(1 - j\lambda \frac{T}{2}\right)}{j\lambda} + k \sum_{i=1}^n \frac{C_i \left(1 - j\lambda \frac{T}{2}\right)}{j\lambda \frac{T}{2} \operatorname{cth} \frac{T}{2T_i}} \quad (6.47)$$

Непосредственное использование подобных выражений для построения соответствующих логарифмических частотных характеристик представляет достаточно сложную задачу. При этом теряется одно из основных достоинств метода логарифмических частотных характеристик — простота построения самих характеристик.

Во многих случаях построение можно значительно упростить, если его производить приближенно и отдельно для областей низких $\lambda = \omega < \frac{2}{T}$ и высоких $\lambda > \frac{2}{T}$ частот [81].

Допустим, что среди n постоянных времени T_i в выражении (6.47) имеется l постоянных времени T_i , удовлетворяющих условию $T_i > \frac{T}{2}$. В соответствующих слагаемых правой части этого выражения можно положить $\frac{2}{T} \operatorname{cth} \frac{T}{2T_i} \approx T_i$; $i = 1, 2, \dots, l$. Тогда для этой области частот (низкие

частоты) можно записать из (6.47) приближенное выражение комплексного коэффициента передачи:

$$W_n(j\lambda) \approx \left(1 - j\lambda \frac{T}{2}\right) \left(\frac{k}{j\lambda} + k \sum_{i=1}^l \frac{C_i}{1 + j\lambda T_i}\right).$$

Из рассмотренных выше примеров (в том числе примера 5 в § 6.3) нетрудно заметить, что это выражение может быть получено непосредственно из передаточной функции непрерывной части подстановкой $s = j\lambda$ и умножением на дополнительный множитель $\left(1 - j\lambda \frac{T}{2}\right)$, т.е.

$$W_n(j\lambda) \approx W_n(s) \Big|_{s=j\lambda} \left(1 - j\lambda \frac{T}{2}\right). \quad (6.48)$$

Псевдочастота λ в этой области практически совпадает с частотой ω входного сигнала.

В выражении (6.48) $W_n(s)$ так называемая низкочастотная часть передаточной функции $W(s)$, в которой присутствуют только постоянные времени τ_j , T_j , удовлетворяющие условиям: $\frac{1}{\tau_j} < \frac{2}{T}$; $\frac{1}{T_j} < \frac{2}{T}$.

Нетрудно заметить, что логарифмические частотные характеристики импульсной системы в области низких частот, т.е. слева от частоты $\lambda = \frac{2}{T}$, будут совпадать с характеристиками непрерывной части.

Для построения логарифмических частотных характеристик в области высоких частот необходимо записать передаточную функцию $W(s)$ для области частот $\omega > \frac{2}{T}$ и, последовательно применив z - и w -преобразования, найти $W_n(j\lambda)$.

Например, если для области частот правее частоты $\omega = \frac{2}{T}$ передаточная функция непрерывной части имеет вид

$$W_B(s) = \frac{k_B}{s(1 + T_{l+1}s)(1 + T_{l+2}s) \dots (1 + T_ns)},$$

где $k_B = \omega_{cp}$ (частота среза непрерывной части системы), то импульсная передаточная функция в области высоких частот имеет вид

$$W_B(z) = \frac{k_B T}{z-1} - k_B \sum_{i=l+1}^n \frac{C_i(1-d_i)}{z-d_i}, \quad d_i = e^{-\frac{T}{T_i}}.$$

Применив w -преобразование с последующей заменой $w = j\lambda \frac{T}{2}$, получим выражение для частотной передаточной функции

$$W_B(j\lambda) = \frac{k_B \left(1 - j\lambda \frac{T}{2}\right) \left[1 + j\lambda \left(\frac{T}{2} - T_\Sigma\right)\right]}{j\lambda \left(1 + j\lambda \frac{T}{2}\right)}, \quad T_\Sigma = \sum_{i=1}^n T_i. \quad (6.49)$$

Объединив выражения (6.48) и (6.49) для низких и высоких частот, получим комплексный коэффициент передачи импульсной системы:

$$W(j\lambda) = W_B(s) \Big|_{s=j\lambda} = \frac{\left[1 + j\lambda \left(\frac{T}{2} - T_\Sigma\right)\right] \left(1 - j\lambda \frac{T}{2}\right)}{\left(1 + j\lambda \frac{T}{2}\right)}.$$

Это выражение легко использовать для построения логарифмических частотных характеристик.

6.6. ПРЕДСТАВЛЕНИЕ ДИСКРЕТНЫХ СИСТЕМ В ПРОСТРАНСТВЕ СОСТОЯНИЙ

Переход к описанию дискретной системы в пространстве состояний можно осуществить различными способами. Далее рассматриваются два из них. Первый способ заключается в прямом переходе от разностного уравнения путем введения в него переменных состояния; второй — основан на использовании аналитического решения векторного дифференциального уравнения непрерывной части импульсной системы с экстраполятором нулевого порядка [33], [37], [44].

1. *Запись разностного уравнения импульсной системы в векторной форме.* Рассмотрим некоторый линейный импульсный фильтр (импульсную систему) с входным сигналом $u(nT)$ и выходным $y(nT)$. Пусть импульсная передаточная функция фильтра задана выражением

$$W(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_q z^{-q}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_q z^{-q}}$$

или

$$W(z) = \frac{b_0 z^q + b_1 z^{q-1} + \dots + b_q}{z^q + a_1 z^{q-1} + \dots + a_q} = \frac{B(z)}{A(z)}. \quad (6.50)$$

Соответствующее разностное уравнение имеет вид

$$\begin{aligned} y(n+q) + a_1 y(n+q-1) + \dots + a_q y(n) = \\ = b_0 u(n+q) + b_1 u(n+q-1) + \dots + b_q u(n). \end{aligned} \quad (6.51)$$

Введем следующие переменные состояния:

$$\begin{aligned} x_1(n) &= y(n); \\ x_2(n) &= y(n+1) = x_1(n+1); \\ x_3(n) &= y(n+2) = x_2(n+1); \\ &\vdots \quad \quad \quad \vdots \\ x_q(n) &= y(n+q-1) = x_{q-1}(n+1); \\ y(n+q) &= x_q(n+1). \end{aligned} \quad (6.52)$$

Рассмотрим вначале частный случай, положив в (6.50), (6.51) $b_q \doteq 1$, $b_0 = b_1 = \dots = b_{q-1} = 0$. Подставив (6.52) в разностное уравнение (6.51), получим:

$$y(n+q) = x_q(n+1) = -a_1 x_q(n) - a_2 x_{q-1}(n) - \dots - a_q x_1(n) + u(n). \quad (6.53)$$

Это уравнение можно представить в виде векторного разностного уравнения состояний:

$$\begin{bmatrix} x_1(n+1) \\ x_2(n+1) \\ \vdots \\ x_q(n+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_q & -a_{q-1} & -a_{q-2} & \dots & -a_1 \end{bmatrix} \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_q(n) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} u(n) \quad (6.54)$$

и уравнения выхода

$$y(n) = [1 \ 0 \ \dots \ 0] \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_q(n) \end{bmatrix}. \quad (6.55)$$

Введем обозначения: x — вектор переменных состояния; A — матрица состояний; B — матрица входа; C — матрица выхода. Тогда можно записать (6.54), (6.55) в виде

$$x(n+1) = Ax(n) + Bu(n); \quad y(n) = Cx(n).$$

Для рассматриваемого частного случая, когда в выражении (6.50) полином $B(z) = 1$, можно записать:

$$Y(z) = \frac{1}{A(z)} U(z) = \frac{1}{z^q + a_1 z^{q-1} + \dots + a_q} U(z). \quad (6.56)$$

Поскольку переменная состояния $x_1(n) = y(n)$, а следовательно, $Y(z) = X_1(z)$, то выражению (6.56) соответствует разностное уравнение вида

$$a_q x_1(n) + a_{q-1} x_1(n+1) + \dots + a_1 x_1(n+q-1) + x_1(n+q) = u(n). \quad (6.57)$$

Если же $b_q \neq 1$, а $b_0, b_1, b_2, \dots, b_{q-1} \neq 0$, то из выражения (6.50) с учетом (6.56) можно записать:

$$Y(z) = B(z)X_1(z) = (b_q + b_{q-1}z + \dots + b_0 z^q) X_1(z),$$

чему соответствует разностное уравнение

$$y(n) = b_q x_1(n) + b_{q-1} x_1(n+1) + \dots + b_0 x_1(n+q).$$

В этом выражении $x_1(n)$ есть решение разностного уравнения (6.57). Используя обозначения, введенные в (6.52), перепишем правую часть последнего уравнения в виде

$$y(n) = b_q x_1(n) + b_{q-1} x_2(n) + \dots + b_1 x_q(n) + b_0 x_q(n+1). \quad (6.58)$$

Подставив в (6.58) выражение (6.53) для $x_q(n+1)$, получим

$$y(n) = (b_q - b_0 a_q) x_1(n) + (b_{q-1} - b_0 a_{q-1}) x_2(n) + \dots + (b_1 - b_0 a_1) x_q(n) + b_0 u(n).$$

Это уравнение можно записать как уравнение выхода в матричной форме

$$y(n) = \left[(b_q - b_0 a_q) \dots (b_1 - b_0 a_1) \right] \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_q(n) \end{bmatrix} + b_0 u(n)$$

или в векторной

$$y(n) = Cx(n) + b_0 u(n).$$

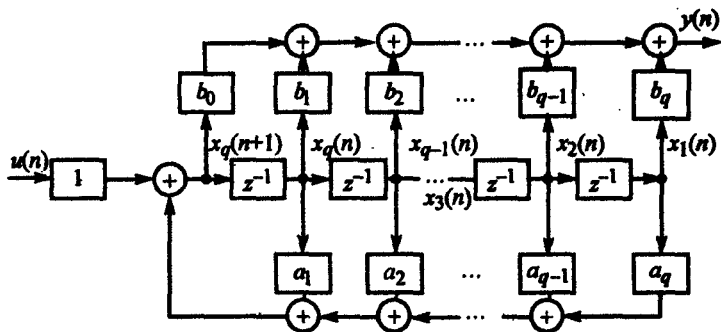


Рис. 6.15. Структурная схема в переменных состояния

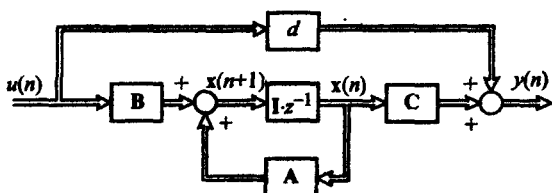


Рис. 6.16. Обобщенная блок-схема в переменных состояния

При $b_0 = 0$ последние два уравнения приобретают вид

$$y(n) = [b_q \dots b_1] \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_q(n) \end{bmatrix};$$

$$y(n) = Cx(n).$$

Структурная схема, соответствующая разностному уравнению, полученному исходя из уравнений (6.53) и (6.58) и записанному в пространстве состояний, представлена на рис. 6.15.

Таким образом, уравнение состояния дискретной системы и уравнение выхода имеют вид

$$x(n+1) = Ax(n) + Bu(n); \quad y(n) = Cx(n) + du(n).$$

Данным уравнениям соответствует блок-схема, показанная на рис. 6.16.

2. Запись разностного уравнения в векторной форме, основанная на решении векторного дифференциального уравнения непрерывной части системы с экстраполятором нулевого порядка. Рас-

смотрим векторное дифференциальное уравнение непрерывной части импульсной системы

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t), \quad (6.59)$$

дополненное уравнением выхода $y(t) = \mathbf{C}\mathbf{x}(t) + du(t)$.

Решение такого дифференциального уравнения при начальном состоянии $\mathbf{x}(0)$ описывается выражением

$$\mathbf{x}(t) = \Phi(t)\mathbf{x}(0) + \int_0^t \Phi(t-\tau)\mathbf{B}u(\tau)d\tau. \quad (6.60)$$

Переходная матрица состояния

$$\Phi(t) = e^{\mathbf{A}t}$$

является матричной экспонентой, которая определяется как ряд (2.38):

$$e^{\mathbf{A}t} = \sum_{v=0}^{\infty} \frac{(\mathbf{A}t)^v}{v!}.$$

Если входной $u(t)$ и выходной $y(t)$ сигналы квантуются по времени, то описание импульсной системы в пространстве состояний можно получить непосредственно из уравнений (6.59), (6.60), полагая, что на входе непрерывной части стоит экстраполятор нулевого порядка, т.е.

$$u(t) = u(nT) \text{ при } nT \leq t < (n+1)T.$$

В этом случае для начального состояния $\mathbf{x}(nT)$ решение (6.60) уравнения состояния на интервале $nT \leq t < (n+1)T$ приобретает вид

$$\mathbf{x}(t) = \Phi(t-nT)\mathbf{x}(nT) + u(nT) \int_{nT}^t \Phi(t-\tau)\mathbf{B}d\tau.$$

Тогда для момента $t = (n+1)T$ состояние определяется как

$$\mathbf{x}[(n+1)T] = \Phi(T)\mathbf{x}(nT) + u(nT) \int_{nT}^{(n+1)T} \Phi[(n+1)T-\tau]\mathbf{B}d\tau.$$

Введя в это уравнение новую переменную $m = (n+1)T - \tau$, для которой $dm = -d\tau$, получим

$$\mathbf{x}(n+1) = \Phi(T)\mathbf{x}(n) + u(n) \int_0^T \Phi(m)\mathbf{B}dm.$$

С целью упрощения записи введем следующие обозначения:

$$\mathbf{A}_1 = \Phi(T) = e^{\mathbf{A}T};$$

$$\mathbf{B}_1 = \int_0^T \Phi(m) \mathbf{B} dm.$$

При этом векторное разностное уравнение состояния принимает вид

$$\mathbf{x}(n+1) = \mathbf{A}_1 \mathbf{x}(n) + \mathbf{B}_1 u(n), \quad (6.61)$$

а уравнение выхода

$$y(n) = \mathbf{C} \mathbf{x}(n) + du(n). \quad (6.62)$$

Применяя линейное неособое преобразование вида

$$\mathbf{v} = \mathbf{T} \mathbf{x},$$

можно изменить базис пространства состояний, т.е. получить другие формы представления импульсных систем. Вместо уравнений (6.61), (6.62) будем иметь уравнения:

$$\mathbf{v}(n+1) = \mathbf{A}_v \mathbf{v}(n) + \mathbf{B}_v u(n);$$

$$y(n) = \mathbf{C}_v \mathbf{v}(n) + du(n),$$

где

$$\mathbf{A}_v = \mathbf{T} \mathbf{A}_1 \mathbf{T}^{-1}, \quad \mathbf{B}_v = \mathbf{T} \mathbf{B}_1, \quad \mathbf{C}_v = \mathbf{C} \mathbf{T}^{-1}.$$

Рассмотрим некоторые способы решения разностных уравнений состояния на примере уравнений (6.61) и (6.62).

Первый из них заключается в построении рекуррентной процедуры решения для заданной последовательности входного сигнала $u(n)$ при начальных условиях $\mathbf{x}(0)$. Эта процедура описывается соотношениями:

$$\mathbf{x}(1) = \mathbf{A}_1 \mathbf{x}(0) + \mathbf{B}_1 u(0);$$

$$\mathbf{x}(2) = \mathbf{A}_1 \mathbf{x}(1) + \mathbf{B}_1 u(1) = \mathbf{A}_1^2 \mathbf{x}(0) + \mathbf{A}_1 \mathbf{B}_1 u(0) + \mathbf{B}_1 u(1);$$

$$\vdots$$

$$\mathbf{x}(n) = \mathbf{A}_1^n \mathbf{x}(0) + \sum_{i=1}^n \mathbf{A}_1^{i-1} \mathbf{B}_1 u(n-i). \quad (6.63)$$

В выражениях (6.63) первый член представляет решение однородного разностного уравнения, а второй — частное решение неоднородного уравнения. Выходной сигнал $y(n)$ определяется из уравнения (6.62).

Второй способ решения используется, если $u(n)$ задано своим z -преобразованием. В этом случае уравнение состояния (6.61) приводится к виду

$$z[\mathbf{X}(z) - \mathbf{x}(0)] = \mathbf{A}_1 \mathbf{X}(z) + \mathbf{B}_1 U(z)$$

или

$$\mathbf{X}(z) = [z\mathbf{I} - \mathbf{A}_1]^{-1} z\mathbf{x}(0) + [z\mathbf{I} - \mathbf{A}_1]^{-1} \mathbf{B}_1 U(z). \quad (6.64)$$

Подставляя это выражение в уравнение выхода (6.62), получим

$$Y(z) = \mathbf{C}[z\mathbf{I} - \mathbf{A}_1]^{-1} z\mathbf{x}(0) + [\mathbf{C}[z\mathbf{I} - \mathbf{A}_1]^{-1} \mathbf{B}_1 + d]U(z). \quad (6.65)$$

Сравнивая уравнение (6.64) с уравнением (6.63), получаем:

$$\mathbf{A}_1^n = \mathbf{Z}^{-1} \{ [z\mathbf{I} - \mathbf{A}_1]^{-1} z \}.$$

Используя уравнение (6.65), можно получить импульсную передаточную функцию системы, связывающую z -изображения $Y(z)$ и $U(z)$ при нулевых начальных условиях $\mathbf{x}(0) = \mathbf{0}$:

$$W(z) = \frac{Y(z)}{U(z)} = \mathbf{C}[z\mathbf{I} - \mathbf{A}_1]^{-1} \mathbf{B}_1 + d.$$

Учитывая, что $[z\mathbf{I} - \mathbf{A}_1]^{-1} = [z\mathbf{I} - \mathbf{A}_1]^* A^{-1}(z)$, где $[z\mathbf{I} - \mathbf{A}_1]^*$ — присодиненная матрица, а

$$A(z) = \det [z\mathbf{I} - \mathbf{A}_1] -$$

— определитель характеристической матрицы, импульсную передаточную функцию можно записать так:

$$W(z) = \frac{B(z)}{A(z)} = \frac{\mathbf{C}[z\mathbf{I} - \mathbf{A}_1]^* \mathbf{B}_1 + dA(z)}{A(z)}.$$

Подставив в уравнение выхода (6.62) выражение (6.63) для $\mathbf{x}(n)$ и положив $\mathbf{x}(0) = \mathbf{0}$, получим

$$y(n) = \sum_{i=1}^n \mathbf{C} \mathbf{A}_1^{i-1} \mathbf{B}_1 u(n-i) + du(n).$$

Если входной сигнал импульсного фильтра имеет вид

$$u(n) = \begin{cases} 1, & n = 0; \\ 0, & n > 0, \end{cases}$$

то решетчатая функция $y(n)$ на выходе будет представлять собой его решетчатую функцию веса $w(n)$, т.е.

$$y(n) = w(n) = \begin{cases} d & \text{при } n = 0; \\ \mathbf{C} \mathbf{A}_1^{n-1} \mathbf{B}_1 & \text{при } n = 1, 2, \dots \end{cases}$$

Откуда можно получить выражение, связывающее импульсную переходную функцию (функцию веса) и импульсную передаточную функцию

$$W(z) = \sum_{n=0}^{\infty} w(n)z^{-n} = d + \sum_{n=1}^{\infty} CA_1^{n-1}B_1z^{-n}.$$

Глава 7

АНАЛИЗ ИМПУЛЬСНЫХ СИСТЕМ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

7.1. СТРУКТУРНЫЕ СХЕМЫ И ПЕРЕДАТОЧНЫЕ ФУНКЦИИ ЗАМКНУТЫХ ИМПУЛЬСНЫХ СИСТЕМ

Рассмотрим замкнутую импульсную систему, структурная схема которой показана на рис. 7.1. Она состоит из идеального импульсного элемента с периодом дискретности T , фиксирующей цепи с передаточной функцией $H(s)$ и непрерывной части, представленной передаточными функциями $W_1(s)$ и $W_2(s)$.

Входной сигнал импульсного элемента $e(t)$ определяется задающим воздействием $g(t)$ и выходной переменной системы $y(t)$:

$$e(t) = g(t) - y(t).$$

Перейдя к соответствующим последовательностям мгновенных импульсов, получим

$$e^*(t) = g^*(t) - y^*(t). \quad (7.1)$$

Уравнение (7.1) называют *уравнением ошибки замкнутой импульсной системы относительно оригиналов*. Перейдя в область z -изображений, получим уравнение ошибки в z -форме

$$E(z) = G(z) - Y(z). \quad (7.2)$$

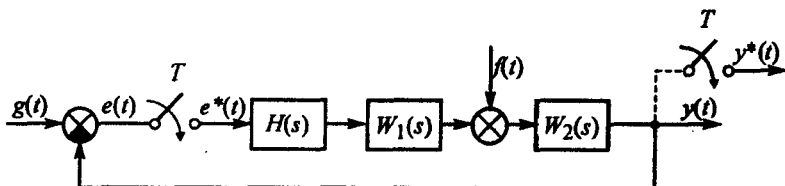


Рис. 7.1. Структурная схема замкнутой импульсной системы

Как показано ранее, уравнение разомкнутой импульсной системы при условии $f(t) \equiv 0$ имеет вид

$$Y(z) = W(z)E(z), \quad (7.3)$$

где

$$W(z) = Z\{H(s)W_1(s)W_2(s)\}.$$

Подставив (7.3) в (7.2), найдем уравнение замкнутой импульсной системы относительно изображения ошибки по задающему воздействию:

$$E_g(z) = \frac{1}{1+W(z)}G(z) = \Phi_{eg}(z)G(z). \quad (7.4)$$

Если подставить (7.4) в (7.3), то получим уравнение замкнутой импульсной системы относительно выходной переменной

$$Y(z) = \frac{W(z)}{1+W(z)}G(z) = \Phi_{yg}(z)G(z).$$

Функция $\Phi_{yg}(z)$ называется импульсной *передаточной функцией замкнутой системы*. Она равна отношению z -изображения выходной переменной к z -изображению сигнала на входе при нулевых начальных условиях:

$$\Phi_{yg}(z) = \frac{Y(z)}{G(z)} = \frac{W(z)}{1+W(z)}.$$

Функцию $\Phi_{eg}(z)$ называют *передаточной функцией замкнутой системы по ошибке*. Она равна отношению z -изображения ошибки к z -изображению воздействия (здесь — задающее воздействие g) при нулевых начальных условиях

$$\Phi_{eg}(z) = \frac{E(z)}{G(z)} = \frac{1}{1+W(z)}.$$

Найдем теперь изображение выходной переменной замкнутой импульсной системы при наличии возмущающего воздействия $f(t)$ и отсутствии задающего воздействия $g(t) \equiv 0$.

При рассмотрении в качестве выходной переменной сигнала $y^*(t)$ удобно преобразовать структурную схему (см. рис. 7.1) к эквивалентному виду, показанному на рис. 7.2.

Выходная переменная $y^*(t)$ такой системы: $y^*(t) = y_1^*(t) + y_2^*(t)$ является выходом фиктивного импульсного элемента на рис. 7.1. Соответствующее уравнение в изображениях имеет вид

$$Y(z) = Y_1(z) + Y_2(z), \quad (7.5)$$

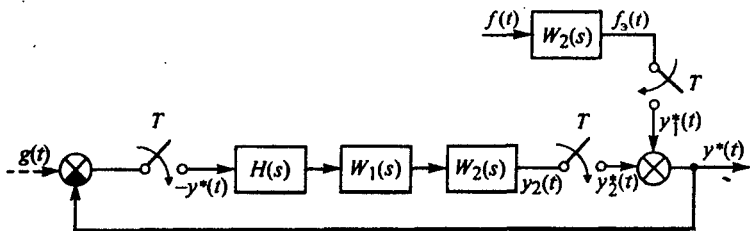


Рис. 7.2. Преобразованная структурная схема замкнутой импульсной системы

где

$$Y_1(z) = Z\{F(s)W_2(s)\} = FW_2(z);$$

$$Y_2(z) = -Y(z)Z\{H(s)W_1(s)W_2(s)\} = -Y(z)W(z).$$

Подставив выражения для $Y_1(z)$, $Y_2(z)$ в уравнение (7.5), получим

$$Y(z) = FW_2(z) - Y(z)W(z).$$

Решим это уравнение относительно изображения выходной переменной $Y(z)$:

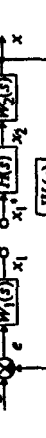
$$Y(z) = \frac{FW_2(z)}{1 + W(z)}.$$

Как видно, выделить в этом выражении импульсную передаточную функцию системы по возмущающемуся воздействию $f(t)$ не представляется возможным. Это вызвано тем, что воздействие приложено не ко входу импульсного элемента. Во всех подобных случаях импульсную передаточную функцию можно записать только для эквивалентного воздействия, полученного пересчетом реального воздействия на вход импульсного элемента.

Для рассматриваемой системы таким воздействием является $f_3(t) = L^{-1}\{F(s)W_2(s)\}$, а соответствующая передаточная функция определяется выражением

$$\Phi_{y_3}(z) = \frac{Y(z)}{F_3(z)} = \frac{1}{1 + W(z)}.$$

Рассмотренный подход к определению импульсных передаточных функций можно использовать для импульсных систем, имеющих и более сложные структуры. В табл. 7.1 [35] показаны примеры таких струк-

Система	Изображение выхода	
	z-изображение $X(z)$	Модифицированное z-изображение $X(z, \sigma)$
	$Z\{H(s)W(s)\}G(z)$	$Z_{\sigma}\{H(s)W(s)\}G(z)$
	$Z\{G(s)W(s)\}$	$Z\{G(s)W(s)\}$
	$\frac{Z\{H(s)W_1(s)\}G(z)}{1 + Z\{H(s)W_1(s)W_2(s)\}}$	$\frac{Z_{\sigma}\{H(s)W_1(s)\}G(z)}{1 + Z\{H(s)W_1(s)W_2(s)\}}$
	$\frac{Z\{H(s)W_2(s)\}Z\{G(s)W_1(s)\}}{1 + Z\{H(s)W_2(s)W_3(s)W_1(s)\}}$	$\frac{Z_{\sigma}\{H(s)W_2(s)\}Z\{G(s)W_1(s)\}}{1 + Z\{H(s)W_2(s)W_3(s)W_1(s)\}}$
	$\frac{D(z)Z\{H(s)W_1(s)\}G(z)}{1 + D(z)Z\{H(s)W_1(s)W_2(s)\}}$	$\frac{Z_{\sigma}\{H(s)W_1(s)\}D(z)G(z)}{1 + D(z)\{H(s)W_1(s)W_2(s)\}}$
	$\frac{Z\{H(s)W_1(s)\}G(z)}{1 + Z\{H(s)W_1(s)\}Z\{H(s)W_2(s)\}}$	$\frac{Z_{\sigma}\{H(s)W_1(s)\}G(z)}{1 + Z\{H(s)W_1(s)\}Z\{H(s)W_2(s)\}}$
	$\frac{D(z)Z\{H(s)W_1(s)\}G(z)}{1 + D(z)Z\{H(s)W_1(s)\}Z\{H(s)W_2(s)\}}$	$\frac{D(z)Z_{\sigma}\{H(s)W_1(s)\}G(z)}{1 + D(z)Z\{H(s)W_1(s)\}Z\{H(s)W_2(s)\}}$
	$\frac{Z\{G(s)W_1(s)\}}{1 + Z\{H(s)W_1(s)W_2(s)\}}$	$\frac{Z_{\sigma}\{G(s)W_1(s)\}}{1 + Z\{H(s)W_1(s)W_2(s)\}}$
	$\frac{Z\{G(s)W_1(s)\}}{1 + Z\{H(s)W_1(s)W_2(s)\}}$	$\frac{Z_{\sigma}\{G(s)W_1(s)\}}{1 + Z\{H(s)W_1(s)W_2(s)\}}$

тур с импульсными элементами, работающими синхронно, синфазно и с одинаковой частотой. Там же приведены выражения для z -изображений выходной переменной каждой системы.

7.2. ПРОЦЕССЫ В ИМПУЛЬСНЫХ СИСТЕМАХ

Как было показано, изображение выходной переменной импульсной системы относительно задающего воздействия имеет вид

$$Y(z) = \Phi(z)G(z), \quad (7.6)$$

причем, импульсная передаточная функция $\Phi(z)$ и изображение воздействия $G(z)$ представляют собой дробно-рациональные функции переменной z :

$$\Phi(z) = \frac{B(z)}{A(z)}, \quad (7.7)$$

$$G(z) = \frac{B_g(z)}{A_g(z)}. \quad (7.8)$$

Значения выходной переменной $y(t)$ импульсной системы в дискретные моменты времени $t = nT$ могут быть найдены с помощью обратного z -преобразования:

$$y(nT) = Z^{-1} \{ \Phi(z)G(z) \}.$$

Известны несколько способов определения оригинала по z -изображению. Рассмотрим два из них, которые наиболее широко распространены в инженерной практике.

1. *Использование формулы обращения и теоремы Коши (A.L. Cauchy) о вычетах.* Искомое решение — переменная $y(nT)$ — может быть получено с использованием так называемой формулы обращения

$$y(nT) = \frac{1}{2\pi j} \oint_{\Gamma} Y(z) z^{n-1} dz. \quad (7.9)$$

Здесь Γ — контур интегрирования плоскости z , охватывающий все полюсы подынтегрального выражения $Y(z)z^{n-1}$. Наиболее часто в качестве контура Γ принимается окружность радиуса R с центром в начале координат плоскости z (рис. 7.3). Радиус R выбирается таким, чтобы все полюсы функции $Y(z)z^{n-1}$ находились внутри этой окружности, т.е. $R > \max |z_i|$, где $z_i (i = 1, 2, \dots)$ — корни уравнения $A(z)A_g(z) = 0$ или особые точки функции $Y(z)$.

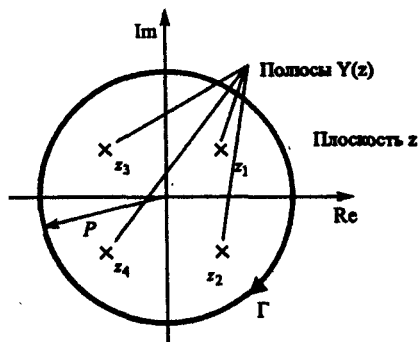


Рис. 7.3. К вычислению контурного интеграла по формуле 7.9

Контурный интеграл (7.9) может быть вычислен с помощью теоремы о вычетах, согласно которой значение интеграла равно сумме всех вычетов подынтегральной функции внутри контура Γ , т. е.

$$y(nT) = \sum_{i=1}^N \text{Res}_{z_i} Y(z) z^{n-1}. \quad (7.10)$$

Здесь $z_i (i = 1, 2, \dots, N)$ — особые точки функции $Y(z)z^{n-1}$. Если изображение $Y(z)$ не содержит множителя z , то число особых точек

$$N = \begin{cases} q+1 & \text{при } n=0; \\ q & \text{при } n > 0, \end{cases}$$

где q — число особых точек функции $Y(z)$.

В технических приложениях изображение $Y(z)$ обычно представляет собой дробно-рациональную функцию z :

$$Y(z) = \frac{B_y(z)}{A_y(z)},$$

особыми точками которой являются ее полюсы $z = z_i$, т.е. корни уравнения $A_y(z) = 0$. Если $Y(z)$ имеет полюс $z = z_i$ кратности μ , то значение вычета в точке $z = z_i$ определяется выражением [81]

$$\text{Res}_{z_i} Y(z) z^{n-1} = \lim_{z \rightarrow z_i} \frac{1}{(\mu-1)!} \frac{d^{\mu-1}}{dz^{\mu-1}} \left[(z - z_i)^\mu Y(z) z^{n-1} \right].$$

При $\mu = 1$ (простой полюс z_i) получается более простая формула для вычета:

$$\text{Res}_{z_i} Y(z) z^{n-1} = \lim_{z \rightarrow z_i} \left[(z - z_i) Y(z) z^{n-1} \right].$$

В случае модифицированного z -изображения в формулу (7.10) вместо $Y(z)$ подставляется $Y(z, \sigma)$.

Также как и в непрерывных системах, выражение для выходной координаты $y(nT)$ при произвольном входном воздействии может быть представлено в виде суммы переходной и установившейся составляющих:

$$Y(nT) = y_{\text{пер}}(nT) + y_{\text{уст}}(nT).$$

Воспользовавшись формулой (7.10), нетрудно получить выражение для этих составляющих движения:

$$y_{\text{пер}}(nT) = \sum_{i=1}^q \text{Res}_{z_i} \Phi(z) G(z) z^{n-1}, \quad (7.11)$$

$$y_{\text{уст}}(nT) = \sum_{j=1}^v \text{Res}_{z_j} \Phi(z) G(z) z^{n-1}, \quad (7.12)$$

где $z_i; i = 1, 2, \dots, q$ — полюсы передаточной функции $\Phi(z)$; $z_j; j = 1, 2, \dots, v$ — полюсы изображения воздействия $G(z)$.

В том случае, когда полюсы функций (7.7) и (7.8) простые, не равные нулю выражения (7.11) и (7.12) принимают вид

$$y_{\text{пер}}(nT) = \sum_{i=1}^q \frac{B(z_i)}{A'(z_i)} G(z_i) z_i^n, \quad (7.13)$$

$$y_{\text{уст}}(nT) = \sum_{j=1}^v \Phi(z_j) \frac{B_g(z_j)}{A_g'(z_j)} z_j^n. \quad (7.14)$$

Как видно из (7.13), (7.14), характер переходной и установившейся составляющих зависит, соответственно, от полюсов импульсной передаточной функции системы и полюсов изображения воздействия. Выражение (7.14) показывает, что для нахождения установившейся составляющей нет необходимости определять полюсы передаточной функции системы. Нахождение же полюсов z_j изображения воздействия для типовых временных функций не составляет труда. Например, если $g(t) = 1(t)$, то изображение $G(z) = z/z - 1$ имеет единственный полюс $z_j = 1$. Из выражения (7.14) сразу получаем установившуюся составляющую выхода системы при ступенчатом воздействии $y_{\text{уст}}(nT) = \Phi(z)|_{z=1}$, которая постоянна при всех n .

2. *Определение оригинала методом разложения функции $Y(z)$ в степенной ряд.* Здесь задача решается путем разложения z -изображения $Y(z)$ в ряд по степеням z^{-1} :

$$Y(z) = \frac{B_y(z)}{A_y(z)} = c_0 + c_1 z^{-1} + c_2 z^{-2} + \dots \quad (7.15)$$

Разложение можно делать любым способом, так как оно единственно. По определению, z -преобразование искомой решетчатой функции $y(nT)$

$$Y(z) = \sum_{n=0}^{\infty} y(nT) z^{-n} = y(0) + y(T) z^{-1} + y(2T) z^{-2} + \dots \quad (7.16)$$

Сравнивая правые части выражений (7.15) и (7.16), замечаем, что $y(nT) = c_n$.

Обычно изображение $Y(z)$ представляет собой рациональную дробь:

$$Y(z) = \frac{b_0 z^m + b_1 z^{m-1} + \dots + b_m}{z^q + a_1 z^{q-1} + \dots + a_q}; \quad m \leq q, \quad (7.17)$$

разложение которой удобно находить делением полинома $B_y(z)$ на полином $A_y(z)$. В результате получим следующие формулы для коэффициентов ряда (7.15):

$$c_0 = b_0;$$

$$c_1 = b_1 - a_1 c_0;$$

$$c_2 = b_2 - a_1 c_1 - a_2 c_0;$$

$$\dots\dots\dots;$$

$$c_k = b_k - a_1 c_{k-1} - a_2 c_{k-2} - \dots - a_{k-1} c_1 - a_k c_0.$$

При $g(t) = \delta(t)$ изображение $G(z) = 1$ и, как следует из (7.12),

$$Y(z) = \Phi(z) = \sum_{n=0}^{\infty} w(nT) z^{-n},$$

т.е. коэффициенты c_n ; $n = 0, 1, 2, \dots$ определяют в этом случае значения решетчатой функции веса в моменты $t = nT$.

В отличие от линейных непрерывных систем, в которых переходный процесс затухает бесконечно долго, в импульсных системах возможны такие условия, при которых этот процесс заканчивается за конечное время, т.е. начиная с некоторого момента времени $n \geq n_0$ временная характеристика равна нулю. Очевидно, что это может быть, когда дробь, стоящая в правой части выражения (7.17), превращается в многочлен, т.е. при $a_1 = a_2 = \dots = a_q = 0$.

Отсюда получаем условие конечной длительности переходных процессов: $A_y(z) = z^q$; $q \geq m$.

В заключение отметим, что помимо рассмотренных двух способов нахождения обратного z -преобразования такую задачу в ряде случаев удобно решать разложением функции $Y(z)$ на простые дроби, имеющие табличную форму (см., например, табл.6.1). Последующий переход к оригиналу не представляет трудностей.

7.3. ОЦЕНКА ТОЧНОСТИ ИМПУЛЬСНЫХ СИСТЕМ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ В УСТАНОВИВШЕМСЯ РЕЖИМЕ

Для оценки точности импульсных автоматических систем в установившемся режиме используют величину установившейся ошибки при различных типовых воздействиях, наиболее характерных для рассматриваемой системы.

В замкнутой импульсной системе (см. рис.7.1) ошибка e , задающее g и возмущающее f воздействия связаны следующим уравнением относительно z -изображений:

$$E(z) = \frac{1}{1+W(z)} G(z) + \frac{FW_2(z)}{1+W(z)} = E_g(z) + E_f(z).$$

Это выражение содержит две составляющие ошибки, первая из которых $E_g(z)$ обусловлена задающим воздействием, а вторая $E_f(z)$ — возмущающим.

Установившаяся ошибка импульсной системы может быть вычислена по выражению, определяющему конечное значение оригинала (см. § 6.2), т.е.

$$e(nT) = \lim_{n \rightarrow \infty} \frac{z-1}{z} E_g(z) + \lim_{n \rightarrow \infty} \frac{z-1}{z} E_f(z).$$

Определим установившуюся ошибку системы по задающему воздействию, положив $f(t) \equiv 0$. Получим

$$e(nT) = e_g(nT) = \lim_{n \rightarrow \infty} \left[\frac{z-1}{z} \frac{1}{1+W(z)} G(z) \right]. \quad (7.18)$$

Если на вход подается постоянное воздействие $g(t) = g_0 1(t)$, z -изображение которого $G(z) = \frac{g_0 z}{z-1}$, то в соответствии с (7.18) установившаяся ошибка системы по положению

$$e(nT) = \lim_{n \rightarrow \infty} \frac{g_0}{z \rightarrow 1} \frac{1}{1+W(z)}.$$

При входном воздействии $g(t) = g_1 t$, линейно зависящем от времени, z -изображение

$$G(z) = \frac{g_1 T z}{(z-1)^2},$$

а установившаяся ошибка, согласно (7.18), определяется выражением

$$e(nT) = \lim_{n \rightarrow \infty} \frac{g_1 T}{z \rightarrow 1} \frac{1}{(z-1)(1+W(z))}$$

и называется *ошибкой системы по скорости*.

Если входной сигнал изменяется с постоянным ускорением, т.е. $g(t) = g_2 t^2/2$, то z -изображение имеет вид

$$G(z) = \frac{T^2}{2} \frac{g_2 z(z+1)}{(z-1)^3}.$$

Установившаяся ошибка

$$e(nT) = \lim_{n \rightarrow \infty} \frac{g_2 T^2}{z \rightarrow 1} \frac{1}{2(z-1)^2(1+W(z))}$$

и называется *ошибкой системы по ускорению*.

В общем случае, если воздействие представлено в виде степенной

функции времени $g(t) = \frac{g_k t^k}{k!}$, его z -изображение можно записать так

$$G(z) = \frac{g_k T^k z P_k(z)}{(z-1)^{k+1}},$$

где $P_k(z)$ — полином степени $(k-1)$ относительно z и такой, что $P_k(1) = 1$.

Установившаяся ошибка по этому воздействию определяется выражением

$$e(nT) = \lim_{n \rightarrow \infty} \frac{g_k T^k}{z \rightarrow 1} \frac{1}{(z-1)^k} \frac{1}{1+W(z)}. \quad (7.19)$$

Допустим, что передаточная функция $W(z)$ разомкнутой системы может быть представлена в виде

$$W(z) = \frac{1}{W_1(z)}, \quad (7.20)$$

где $\nu = 0, 1, 2, \dots$ — порядок астатизма системы, а $W_1(z)$ не имеет нулей и полюсов, равных единице. Подставив (7.20) в выражение (7.19), после несложных преобразований получим:

$$e(nT) = \begin{cases} 0 & \text{при } k < \nu; \\ \frac{g_k T^k}{W_1(1)} & \text{при } k = \nu; \\ \infty & \text{при } k > \nu. \end{cases}$$

Таким образом, для того чтобы импульсная система имела нулевую установившуюся ошибку по задающему воздействию, необходимо чтобы порядок ее астатизма ν превышал степень входного воздействия k .

Рассмотренный способ определения установившейся ошибки системы удобен только при типовых входных воздействиях. Если функция $g(t)$ имеет произвольную форму, достаточно плавную вдали от начальной точки, т.е. имеет там конечное число производных, не равных нулю, то для вычисления ошибки можно воспользоваться понятием коэффициентов ошибок. Разложив передаточную функцию $\Phi_e(z)$ системы по сигналу ошибки для задающего воздействия в степенной ряд по $(1 - z^{-1})$, получим

$$\begin{aligned} \Phi_e(z) = \frac{1}{1+W(z)} = C_0 + \frac{C_1}{T}(1-z^{-1}) + \frac{C_2}{2!T^2}(1-z^{-1})^2 + \dots + \\ + \frac{C_m}{m!T^m}(1-z^{-1})^m + \dots \end{aligned} \quad (7.21)$$

Коэффициенты $C_0, C_1, C_2, \dots, C_m, \dots$ называют *коэффициентами ошибок*. Из соотношения (7.21) следует, что для статической системы ($\nu = 0$)

$$C_0 = \lim_{z \rightarrow 1} \frac{1}{1+W(z)} = \frac{1}{1+k},$$

где $k = W(1)$ — коэффициент передачи системы.

Для системы с астатизмом первого порядка ($\nu = 1$) $C_0 = 0$, поэтому из (7.21) имеем

$$C_1 = \lim_{z \rightarrow 1} \frac{T}{(1-z^{-1})[1+W(z)]} = \frac{T}{k_\nu},$$

где $k_\nu = W_1(1)$ — добротность системы по скорости.

Для системы с астатизмом второго порядка $C_0 = C_1 = 0$, поэтому из (7.21) следует, что

$$C_2 = \lim_{z \rightarrow 1} \frac{2T^2}{(1-z^{-1})^2 [1+W(z)]} = \frac{2T^2}{k_\epsilon},$$

где $k_\epsilon = W_1(1)$ — добротность системы по ускорению.

Аналогичным образом могут быть найдены и другие коэффициенты ошибок C_m .

По коэффициентам ошибок можно построить ряд, представляющий решетчатую функцию $e(nT)$, начиная с некоторого момента $t = kT$ после начала переходного процесса. Такой ряд имеет вид:

$$e(nT) = C_0 g(nT) + C_1 g'(nT) + \frac{C_2}{2!} g''(nT) + \dots + \frac{C_m}{m!} g^{(m)}(nT) + \dots, \quad (7.22)$$

где $g', g'', \dots, g^{(m)}$ — производные функции $g(t)$.

Следует заметить, что применимость рассмотренного способа расчета установившейся ошибки зависит от скорости сходимости ряда (7.22).

7.4. УСТОЙЧИВОСТЬ ИМПУЛЬСНЫХ СИСТЕМ

Линейная импульсная система считается устойчивой, если ограниченному внешнему воздействию соответствует ограниченная реакция. Условие устойчивости для импульсных систем можно записать в виде

$$\lim_{n \rightarrow \infty} y_{\text{пер}}(nT) = 0. \quad (7.23)$$

Иными словами, импульсная система устойчива, если переходная составляющая движения с течением времени стремится к нулю.

Действительно, как видно из выражения (7.14), при ограниченном воздействии, которому соответствует условие $|z_j| < 1; j = 1, 2, \dots, v$, составляющая $y_{\text{пер}}(nT)$ будет также ограниченной. Для того чтобы полная реакция была ограниченной, необходимо в выражении (7.13) иметь $|z_i| < 1; i = 1, 2, \dots, q$, что и обеспечивает выполнимость условия (7.23).

Таким образом, все корни характеристического полинома $A(z)$ или полюсы передаточной функции $\Phi(z)$ импульсной системы должны быть по модулю меньше единицы, т.е. расположены в плоскости z внутри окружности единичного радиуса с центром в начале координат (рис. 7.4).

Если хотя бы один полюс или пара комплексно-сопряженных полюсов передаточной функции $\Phi(z)$ лежат вне единичного круга ($|z_i| > 1$), то импульсная система неустойчива. Если имеются полюсы на окружности единичного радиуса ($|z_i| = 1$), а все остальные — внутри круга, то

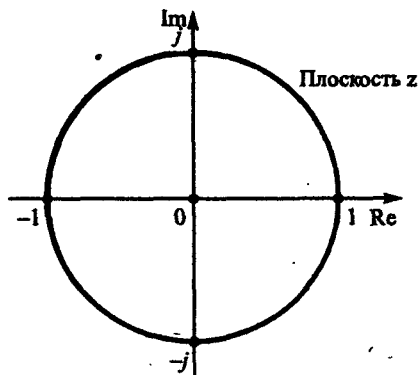


Рис. 7.4. К определению условий устойчивости в плоскости z

импульсная система находится на границе устойчивости. Если это полюс $|z_i| = +1$, то импульсную систему называют *нейтральной*.

Таким образом, исследование устойчивости импульсных систем сводится к определению расположения корней характеристического полинома $A(z)$ относительно окружности единичного радиуса. Для решения этой задачи могут быть применены непосредственно или модифицированы все критерии устойчивости, используемые в теории непрерывных систем.

Так, например, алгебраический критерий Рауса-Гурвица (см. п. 3.3.2), устанавливающий принадлежность всех корней характеристического полинома левой полуплоскости, может быть применен и для импульсных систем с характеристическим полиномом

$$A_w(w) = A(z) \Big|_{z=\frac{1+w}{1-w}} = a'_n w^n + a'_{n-1} w^{n-1} + \dots + a'_1 w + a_0.$$

При использовании критерия Найквиста (см. § 3.5) строят (или снимают экспериментально) частотные характеристики импульсной системы в разомкнутом состоянии. Это могут быть амплитудно-фазовые частотные характеристики (см. § 6.4) или логарифмические характеристики $L_p(\lambda)$, $\varphi_p(\lambda)$ (см. § 6.5).

Формулировки критерия Найквиста для импульсных систем повторяют приведенные в § 3.5 с учетом особенностей частотных характеристик этих систем (например, периодичность по ω). Вместо числа правых полюсов передаточной функции $W_p(s)$ непрерывной системы в критерии Найквиста для импульсных систем учитывается число полюсов $W_p(z)$, модуль которых $|z_i| > 1$.

СИНТЕЗ ИМПУЛЬСНЫХ СИСТЕМ УПРАВЛЕНИЯ**8.1. О СИНТЕЗЕ ИМПУЛЬСНЫХ СИСТЕМ**

Как было показано в гл. 5, дискретные системы управления образованы взаимодействующими непрерывными (объекты, исполнительные механизмы) и дискретными (измерительные преобразователи, управляющие устройства) элементами. Для имитационных исследований таких «смешанных» систем, включающих преобразователи непрерывных и дискретных сигналов (импульсных или числовых последовательностей), разработаны специальные языки и программные средства. Для решения задач синтеза обычно используются однородные математические модели систем — непрерывные либо дискретные.

Существуют два основных подхода к синтезу импульсных систем. В том случае, когда период квантования времени мал, расчеты проводятся по непрерывным моделям. При необходимости цифровой реализации полученного непрерывного алгоритма управления осуществляется его дискретизация (дискретная аппроксимация непрерывного алгоритма). Период квантования выбирается из условия сохранения достигнутого при синтезе качества процессов. Если же в постановке задачи синтеза изначально оговаривается цифровая реализация алгоритма управления с необязательно малым значением периода квантования времени, предусматривается конечная длительность переходных процессов или учитываются другие особенности, присущие дискретным системам, то синтез проводится по дискретным математическим моделям. В процессе проектирования часто оба подхода реализуются многократно для различных значений периода квантования.

Следует отметить, что этап синтеза по требованиям к установившимся процессам в системе слабо зависит от выбора непрерывной или дискретной модели, так как для медленных движений (низких частот) эффект квантования времени проявляется слабо.

Далее основное внимание уделено частотному методу синтеза импульсных систем по требованиям к качеству переходных процессов, когда задача сводится к определению передаточной функции или/и частотной характеристики управляющего устройства (компенсатора, корректора), при введении которых система обладает заданными свойствами [35], [81].

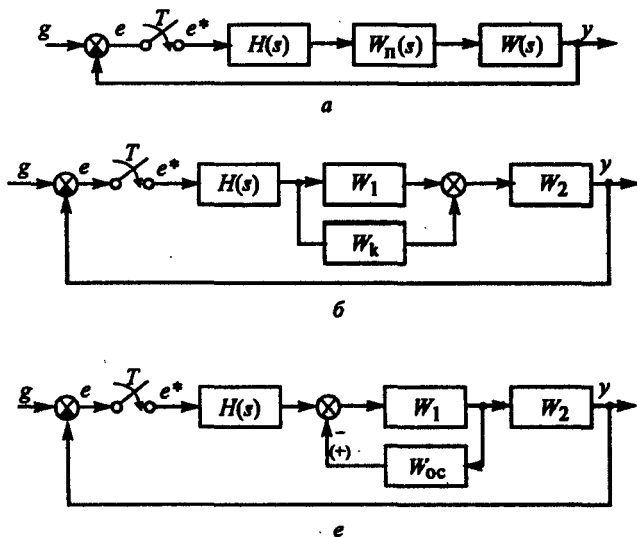


Рис. 8.1. Различные варианты включения непрерывных корректирующих устройств в импульсных системах

Коррекцию импульсных систем можно осуществить с помощью различных устройств, которые по своему действию разделяются на непрерывные и дискретные.

Непрерывная коррекция импульсной системы сводится к коррекции ее непрерывной части введением последовательных (рис. 8.1, а) или параллельных (рис. 8.1, б) корректирующих устройств, либо местной положительной или отрицательной обратной связи (рис. 8.1, в).

При **дискретной коррекции** в качестве корректирующих устройств используются цифровые вычислители или импульсные фильтры. Соответствующие структурные схемы импульсных систем показаны на рис. 8.2.

Импульсные передаточные функции систем (см. рис. 8.2) для разомкнутого состояния имеют следующий вид:

- при последовательной коррекции (рис. 8.2, а)

$$W(z) = \frac{z-1}{z} Z \left\{ \frac{W(s)}{s} \right\} D_n(z);$$

- при параллельной коррекции (рис. 8.2, б)

$$W(z) = \frac{z-1}{z} Z \left\{ \frac{W_1(s)W_2(s)}{s} \right\} + D_k(z) \frac{z-1}{z} Z \left\{ \frac{W_2(s)}{s} \right\};$$

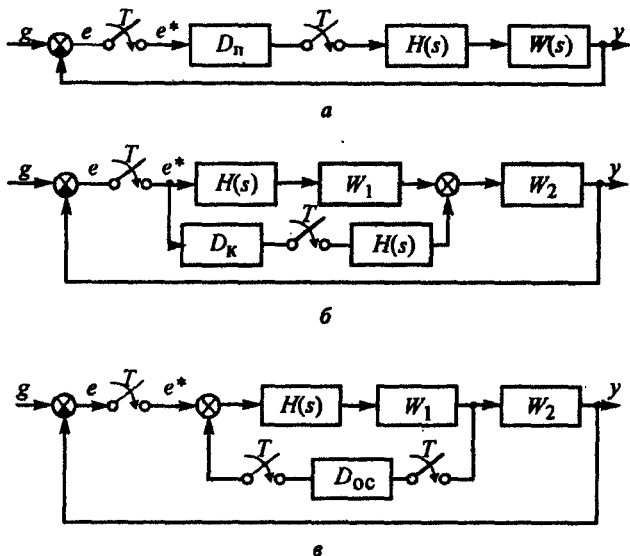


Рис. 8.2. Различные варианты включения дискретных корректирующих устройств

- для коррекции местной отрицательной связью (рис. 8.2, в)

$$W(z) = \frac{\frac{z-1}{z} Z \left\{ \frac{W_1(s)W_2(s)}{s} \right\}}{1 + \frac{z-1}{z} Z \left\{ \frac{W_1(s)}{s} \right\} D_{oc}(z)},$$

где $D_n(z)$, $D_k(z)$ и $D_{oc}(z)$ — импульсные передаточные функции дискретных корректирующих устройств. Приведенные выражения справедливы при использовании формирующего устройства типа экстраполятора нулевого порядка (фиксатора).

Выбор способа дискретной коррекции в каждом отдельном случае направлен на получение наиболее простой реализации дискретного корректирующего устройства.

Важнейшим является этап формализации требований к качеству установившихся и переходных процессов в виде желаемой передаточной функции замкнутой или разомкнутой системы. Ясно, что результат здесь не является однозначным, и представляет собой компромисс между требованиями к процессам и ограничениями в виде динамических свойств объекта, условиями реализуемости алгоритма управления, грубости системы и др.

8.2. ПОСТРОЕНИЕ ЖЕЛАЕМЫХ ЧАСТОТНЫХ ХАРАКТЕРИСТИК

При синтезе импульсных систем с использованием частотного метода прежде всего необходимо определить желаемую логарифмическую частотную характеристику импульсной системы в соответствии с заданными показателями качества [35], [81].

Во многих реальных импульсных системах период квантования T достаточно мал и соблюдается условие $\frac{1}{T} \gg \omega_{\text{ср}}$ ($\omega_{\text{ср}}$ — частота среза системы). В этом случае частотные характеристики импульсной системы в существенном диапазоне частот мало отличаются от частотных характеристик непрерывной части, квантование по времени практически не проявляется, а значит расчет таких импульсных систем полностью повторяет расчет аналогичных непрерывных систем.

Если же период дискретности таков, что величина $\frac{1}{T}$ соизмерима с частотой среза системы $\omega_{\text{ср}}$, то высокочастотная и среднечастотная части частотной характеристики импульсной системы будут сильно отличаться от соответствующих частей частотной характеристики ее непрерывной части. Здесь желаемую частотную характеристику импульсной системы необходимо строить с учетом квантования по времени. В противном случае показатели качества скорректированной импульсной системы будут значительно отличаться от желаемых.

В достаточно общем случае желаемая передаточная функция непрерывной системы имеет вид

$$W_{\text{ж}}(s) = \frac{k \prod_{j=1}^m (1 + \tau_j s)}{s^v \prod_{i=1}^n (1 + T_i s)} \quad (8.1)$$

Для определения соответствующей желаемой передаточной функции $W_{\text{ж}}(z)$ или комплексного коэффициента передачи $W_{\text{ж}}(j\lambda)$ системы с квантованием по времени, которое необходимо учесть, выражение (8.1) представляется в виде суммы простых дробей (см. § 6.3):

$$W_{\text{ж}}(s) = \sum_i F_i(s),$$

каждая из которых подвергается преобразованию:

$$\frac{z-1}{z} Z\{F_i(s)/s\} = F_i(z).$$

Искомое выражение для желаемой передаточной функции $W_{\Sigma}(z)$ будет, очевидно,

$$W_{\Sigma}(z) = \sum_i F_i(z).$$

После подстановки

$$z = \frac{1 + j\lambda \frac{T}{2}}{1 - j\lambda \frac{T}{2}} \quad (8.2)$$

получим соответствующий комплексный коэффициент передачи $W_{\Sigma}(j\lambda)$, определяющий желаемую частотную характеристику $L_{\Sigma}(\lambda)$.

Реально такой путь оказывается чрезвычайно сложным. Решение задачи можно сильно упростить, если построение логарифмической частотной характеристики $L_{\Sigma}(\lambda)$ осуществлять отдельно для области низких $\omega < \frac{2}{T}$ и высоких $\omega > \frac{2}{T}$, частот, т.е. так, как это делалось в § 6.5.

Как известно, обычно параметры желаемой логарифмической характеристики $L_{\Sigma}(\omega)$ выбираются так, чтобы в окрестности частоты среза $\omega_{\text{ср}}$ ее наклон был равен — 20 дБ/дек (см. гл. 4), а все частоты сопряжений $\frac{1}{\tau_j} < \omega_{\text{ср}}$ ($j = 1, 2, \dots, m$).

При синтезе желаемой частотной характеристики импульсной системы, кроме того, практически во всех случаях необходимо выполнить условие $\frac{2}{T} > \omega_{\text{ср}}$, которое согласуется с известной теоремой Котельникова-Шеннона и дает возможность обеспечить необходимые запасы устойчивости системы.

Учитывая изложенное, можно записать выражения для низкочастотной и высокочастотной частей передаточной функции (8.1):

$$W_{\text{н}}(s) = \frac{k(1 + \tau_1 s) \dots (1 + \tau_m s)}{s^v (1 + T_1 s) \dots (1 + T_q s)}; \quad (8.3)$$

$$W_{\text{в}}(s) = \frac{\omega_{\text{ср}}}{s(1 + T_{q+1} s) \dots (1 + T_n s)}, \quad (8.4)$$

где предполагается, что среди постоянных времени T_i имеются большие ($i = 1, 2, \dots, q$), которым соответствуют частоты сопряжений, меньшие ω_{cp} , и малые ($i = q + 1; q + 2, \dots, n$), дающие частоты сопряжений, большие ω_{cp} (правее частоты среза).

Импульсные передаточные функции $W_H(z)$ и $W_B(z)$, соответствующие выражениям (8.3), (8.4), можно записать, используя примеры из § 6.3 (см. (6.37), (6.38) и (6.39)). Эти передаточные функции для $\nu = 1$ имеют вид

$$W_H(z) = \frac{kT}{z-1} + k \sum_{i=1}^q \frac{c_i(1-d_i)}{z-d_i}, \quad d_i = e^{-\frac{T}{T_i}}; \quad (8.5)$$

$$W_B(z) = \frac{\omega_{cp}T}{z-1} + \omega_{cp} \sum_{i=q+1}^n \frac{c_i(1-d_i)}{z-d_i}. \quad (8.6)$$

Для других значений ν будем иметь иные выражения $W_H(z)$ (см., например, (6.39) для $\nu = 2$). Сделав подстановку (8.2) в (8.5), (8.6) и учитывая, что $\frac{T}{2} \operatorname{cth} \frac{T}{2T_i} \approx T_i$ для $i = 1, 2, \dots, q$ и $\frac{T}{2} \operatorname{cth} \frac{T}{2T_i} \approx \frac{T}{2}$ для $i = q + 1, q + 2, \dots, n$, получим приближенные выражения (см. примеры из § 6.3):

$$W_H(j\lambda) \approx W_H(s)|_{s=j\lambda} \left(1 - j\lambda \frac{T}{2} \right);$$

$$W_B(j\lambda) \approx \frac{\omega_{cp} \left(1 - j\lambda \frac{T}{2} \right)}{j\lambda \left(1 + j\lambda \frac{T}{2} \right)} \left[1 + j\lambda \left(\frac{T}{2} - T_\Sigma \right) \right], \quad (8.7)$$

где $T_\Sigma = \sum_{q+1}^n T_i$ — сумма малых постоянных времени. Очевидно, что выражение (8.7) справедливо при любых значениях ν и, даже более того, — при любых передаточных функциях $W_H(s)$.

Поскольку начало логарифмической частотной характеристики $L_H(\lambda)$ сливается («сшивается») с концом частотной характеристики $L_B(\lambda)$ в точке $\lambda_{cp} \approx \omega_{cp}$, можно записать общее приближенное выражение для $W_\Sigma(j\lambda)$ в виде

$$W_\Sigma(j\lambda) = W_H(s)|_{s=j\lambda} \frac{\left(1 - j\lambda \frac{T}{2} \right)}{\left(1 + j\lambda \frac{T}{2} \right)} \left[1 + j\lambda \left(\frac{T}{2} - T_\Sigma \right) \right]. \quad (8.8)$$

Частотная передаточная функция (8.8) задает вид соответствующей желаемой частотной характеристики $L_{\star}(\lambda)$ импульсной системы. Параметры этой характеристики должны выбираться с учетом требований к точности установившихся режимов и к качеству переходных процессов. При этом возможны две постановки задачи:

- 1) период T квантования (дискретности) задан;
- 2) требуется найти максимально допустимое значение T .

Построение желаемой частотной характеристики $L_{\star}(\lambda)$ начинают с построения частотной характеристики $L_{\star}(\omega)$ без учета временной дискретизации (квантования по времени) так, как это было показано в гл. 4. По характеристике $L_{\star}(\omega)$ записывается передаточная функция $W_{\star}(s)$ вида (8.1), а затем — частотная передаточная функция $W_{\star}(j\lambda)$ импульсной системы по выражению (8.8). Очевидно, что типовым частотным характеристикам $L_{\star}(\omega)$ в области средних частот (см. гл. 4) будут соответствовать также типовые характеристики $L_{\star}(\lambda)$. На рис. 8.3, а-в показаны примеры таких характеристик, а в табл. 8.1 приведены соответствующие передаточные функции непрерывной системы и частотные передаточные функции той же системы с квантованием по времени (импульсной системы).

В области низких частот ($\omega \approx \lambda < \lambda_{\text{ср}} \approx \omega_{\text{ср}}$) желаемая логарифмическая частотная характеристика импульсной системы совпадает с желаемой характеристикой соответствующей непрерывной системы; параметры жела-

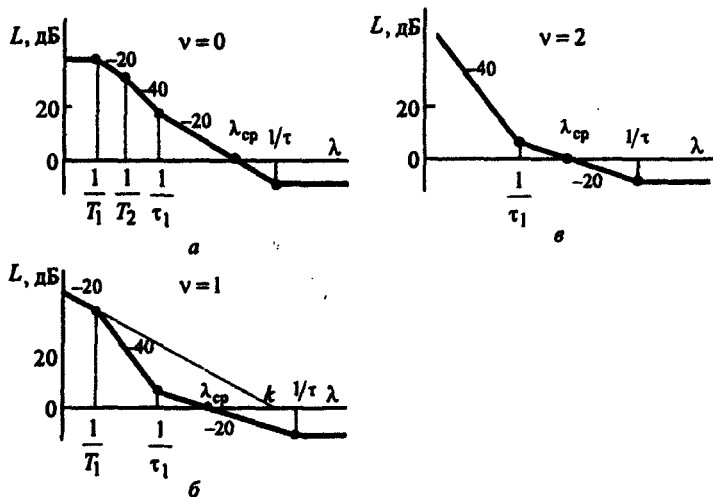


Рис. 8.3. Типовые желаемые логарифмические частотные характеристики импульсных систем

Тип	Степень астатизма	Передающая функция непрерывной части	Частотная характеристика и импульсной системы с фиксатором нулевого порядка
I	0	$\frac{k(1+\tau_1 s)}{(1+T_1 s)(1+T_2 s)\dots(1+T_n s)}$	$k(1+j\lambda\tau_1)\left(1-j\lambda\frac{T}{2}\right)^{k+1} \times$ $\times \left[1+j\lambda\frac{T}{2}(1-2\varepsilon)\right]$ $\frac{T_\Sigma}{T} = k + \varepsilon; \quad T_\Sigma = \tau_s = \sum_{i=3}^n T_i;$ $k = 0, 1, 2, \quad 0 \leq \varepsilon \leq 1$
II	1	$\frac{k(1+\tau_1 s)}{s(1+T_1 s)(1+T_2 s) \times (1+T_3 s)\dots(1+T_n s)}$	$k(1+j\lambda\tau_1)\left(1-j\lambda\frac{T}{2}\right)^{k+1} \times$ $\times \left[1+j\lambda\frac{T}{2}(1-2\varepsilon)\right]$ $\frac{\quad}{j\lambda(1+j\lambda T_1)\left(1+j\lambda\frac{T}{2}\right)^{k+1}},$ $T_\Sigma = \tau_s = \sum_{i=2}^n T_i;$
III	2	$\frac{k(1+\tau_1 s)}{s^2(1+T_1 s) \times (1+T_2 s)\dots(1+T_n s)}$	$k(1+j\lambda\tau_1)\left(1-j\lambda\frac{T}{2}\right)^{k+1} \times$ $\times \left[1+j\lambda\frac{T}{2}(1-2\varepsilon)\right]$ $\frac{\quad}{(j\lambda)^2\left(1+j\lambda\frac{T}{2}\right)^{k+1}},$ $T_\Sigma = \tau_s = \sum_{i=1}^n T_i;$

емых характеристик в этой области частот выбираются исходя из требований, предъявляемых к точности в установившемся режиме (см. гл. 4).

В области средних частот все желаемые характеристики имеют одинаковый наклон (-20 дБ/дек). Их параметры в данной области частот определяют запас устойчивости системы, вид и время затухания переходного процесса. При определении этих параметров можно использовать различные критерии качества, известные в теории линейных непрерывных систем. Одним из таких критериев является показатель колебательности системы [81]

$$M = \max_{0 \leq \lambda < \infty} \frac{|\Phi(j\lambda)|}{|\Phi(0)|},$$

который особенно удобно использовать для систем с колебательными процессами ($1,3 \leq M \leq 2$). Он позволяет получить простые аналитические соотношения, определяющие параметры желаемой логарифмической частотной характеристики.

Для импульсных систем, имеющих передаточные функции вида (8.12) (см. примеры в табл.8.1), заданный показатель колебательности M будет достигаться, если выполнить условия:

1) в области частот, меньших частоты среза,

$$\sum_{j=1}^m \frac{1}{\tau_j} - \sum_{i=1}^q \frac{1}{T_i} \leq \omega_{\text{ср}} \frac{M-1}{M} \quad (8.9)$$

или

$$\sum_{j=1}^m \frac{1}{\tau_j} - \sum_{i=1}^q \frac{1}{T_i} \leq \omega_e \sqrt{\frac{M-1}{M}};$$

2) в области частот, больших частоты среза,

$$\frac{T}{2} + T_{\Sigma} \leq \frac{1}{\omega_{\text{ср}}} \frac{M}{M+1}; \quad (8.10)$$

или

$$\frac{T}{2} + T_{\Sigma} \leq \frac{1}{\omega_e} \sqrt{\frac{M(M-1)}{M+1}},$$

где $\omega_e \approx \lambda_e$ — значение частоты (псевдо частоты) на рис. 8.3, при которой продолжение асимптоты — 40 дБ/дек пересечет ось λ .

Условия (8.9) полностью совпадают с аналогичными условиями для непрерывных систем, а (8.10) — отличаются, так как в области высоких частот характеристики непрерывной и импульсной систем существенно различны.

8.3. СИНТЕЗ НЕПРЕРЫВНЫХ КОРРЕКТИРУЮЩИХ УСТРОЙСТВ

В случае непрерывной коррекции целенаправленно изменяют характеристики непрерывной части импульсной системы. Различные варианты включения непрерывных корректирующих устройств показаны на рис. 8.1.

При расчете передаточных функций таких устройств целесообразно перейти от желаемой логарифмической частотной характеристики импульсной системы $L_{\text{ж}}(\lambda)$ к желаемой характеристике $L_{\text{нх}}(\omega)$ ее непрерывной части. Тогда расчет логарифмических частотных характеристик корректирующих устройств будет полностью совпадать с аналогичным расчетом для непрерывных систем.

В области низких частот ($\omega \approx \lambda < \lambda_{\text{ср}}$) желаемые логарифмические частотные характеристики непрерывной части $L_{\text{ж}}(\omega)$, $\varphi_{\text{ж}}(\omega)$ и импульсной системы $L_{\text{ж}}(\lambda)$, $\varphi_{\text{ж}}(\lambda)$ совпадают. При построении желаемой логарифмической частотной характеристики непрерывной части в области высоких частот ($\omega > \omega_{\text{ср}}$) целесообразно выбирать малые постоянные времени так, чтобы сопрягающие частоты нескорректированной и желаемой частотных характеристик непрерывной части совпадали. В этом случае корректирующая цепь получается наиболее простой. Однако при этом необходимо, чтобы сумма малых постоянных времени T_i в желаемой передаточной функции непрерывной части системы (эти постоянные времена удовлетворяют условию $\frac{1}{T_i} < \omega_{\text{ср}}$) не превышала максимально

допустимое значение, определяемое соотношением (8.10). В противном случае потребуется либо уменьшить период дискретности, либо ввести дополнительные корректирующие цепи для области высоких частот.

Пример. Определим вид и параметры последовательной корректирующей цепи в импульсной следящей системе с АИМ и экстраполятором нулевого порядка, обеспечивающей максимальную скорость слежения $\dot{g}_{\text{max}} = 20$ град/с и максимальное ускорение слежения $\ddot{g}_{\text{max}} = 5$ град/с² при максимально допустимой установившейся ошибке слежения $e_{\text{max}} = 5$ угл. мин. и показателе колебательности $M = 1,5$.

Передаточная функция нескорректированной непрерывной части системы

$$W(s) = \frac{k}{s(1+T_I s)(1+T_{II} s)(1+T_{III} s)}, \quad (8.11)$$

где $T_I = 0,1$ с; $T_{II} = 0,01$ с; $T_{III} = 0,005$ с.

При этом рассмотрим два варианта расчета:

- 1) максимально допустимый период дискретности системы требуется определить;
- 2) период дискретности задан и равен $T = 0,03$ с.

Вариант 1. В качестве желаемой логарифмической частотной характеристики примем характеристику типа II (см. табл. 8.1), обеспечивающую астатизм первого порядка. Ей соответствует частотная передаточная функция вида

$$W_{\Sigma}(j\lambda) = \frac{k(1+j\lambda\tau_1)\left(1-j\lambda\frac{T}{2}\right)\left[1+j\lambda\left(\frac{T}{2}-T_{\Sigma}\right)\right]}{j\lambda(1+j\lambda T_1)\left(1+j\lambda\frac{T}{2}\right)}. \quad (8.12)$$

Исходя из заданной точности в установившемся режиме определим координаты (ω_k, L_k) контрольной точки на плоскости логарифмических характеристик в области низких частот. Для этого используем понятие эквивалентного синусоидального воздействия. Частота этого воздействия

$$\omega_3 = \ddot{g}_{\max} / \dot{g}_{\max} = 0,25 \text{ с}^{-1} = \omega_k.$$

Требуемое значение усиления в контрольной точке (на контрольной частоте ω_k)

$$L_k(\omega_k) = 20 \lg \frac{\dot{g}_{\max}}{\varepsilon_{\max} \ddot{g}_{\max}} = 20 \lg 954 \approx 60 \text{ дБ}.$$

Необходимое значение добротности по скорости

$$k_v = k = \frac{\dot{g}_{\max}}{\varepsilon_{\max}} \sqrt{2} = \frac{20 \cdot 60 \cdot \sqrt{2}}{5} \approx 340 \text{ с}^{-1}.$$

Значение добротности увеличено в $\sqrt{2}$ раз по отношению к минимально допустимому, поскольку желаемая ЛАХ в низкочастотной области формируется так, чтобы она проходила выше контрольной точки на 3 дБ (в линейном масштабе $\sqrt{2}$). Это необходимо, если первая частота сопряжения характеристики L_{Σ} совпадает с частотой ω_k .

Выберем значение максимальной постоянной времени T_1 в частотной передаточной функции (8.12), равное

$$T_1 = \omega_k^{-1} = 4 \text{ с}.$$

Частота ω_e логарифмической характеристики

$$\omega_e = \sqrt{\frac{\dot{g}_{\max}}{\varepsilon_{\max}}} \sqrt{2} = \sqrt{\frac{5 \cdot 60}{5}} \sqrt{2} \approx 9,2 \text{ с}^{-1}.$$

Значение постоянной времени τ_1 получим из условия (8.9):

$$\tau_1 = \frac{1}{\omega_e \sqrt{\frac{M-1}{M} + \frac{1}{T_1}}} = \frac{1}{9,2 \sqrt{\frac{0,5}{1,5} + 0,25}} \approx 0,18 \text{ с}.$$

Частота среза системы

$$\lambda_{\text{ср}} \equiv \omega_{\text{ср}} = k_v \frac{\tau_1}{T_1} = 340 \frac{0,18}{4} \approx 15,3 \text{ с}^{-1}.$$

В области низких частот, где $\lambda \approx \omega$, желаемая ЛАЧХ импульсной системы $L_{\Sigma}(\lambda)$ и желаемая ЛАЧХ ее непрерывной части $L_{\Sigma}(\omega)$ совпадают.

В области высоких частот вид и параметры желаемой логарифмической характеристики непрерывной части $L_{\Sigma}(\omega)$ выбираем так, чтобы ее сопрягающие частоты совпадали с сопрягающими частотами логарифмической характеристики нескорректированной (исходной) непрерывной части $L_{\Sigma}(\omega)$ (рис. 8.4), где $T_2 = T_{II}$; $T_3 = T_{III}$. Характеристика L_{Σ} не должна заходить в заштрихованную область для того, чтобы выполнить требования, предъявляемые к установившейся точности системы.

Подсчитав сумму малых постоянных времени:

$$T_{\Sigma} = 2T_2 + T_3 = 2T_{II} + T_{III} = 0,01 \cdot 2 + 0,005 = 0,025 \text{ с},$$

из условий (8.10) определяем максимально допустимую величину периода дискретности:

$$T \leq \frac{2}{\omega_{\text{ср}}} \frac{M}{M+1} - 2T_{\Sigma} = 0,08 - 0,05 = 0,03 \text{ с}.$$

Определив значение $\varepsilon = \frac{T_{\Sigma}}{T} = 0,83$, найдем последнюю частоту сопряжения желаемой логарифмической частотной характеристики импульсной системы:

$$\lambda_{\Sigma} = \frac{2}{|1 - 2\varepsilon|T} = \frac{2}{0,66 \cdot 0,03} \approx 100 \text{ с}^{-1}.$$

Вычитанием из желаемой логарифмической характеристики непрерывной части $L_{\Sigma}(\omega)$ логарифмической характеристики $L_{\Sigma}(\omega)$ определим логарифмическую

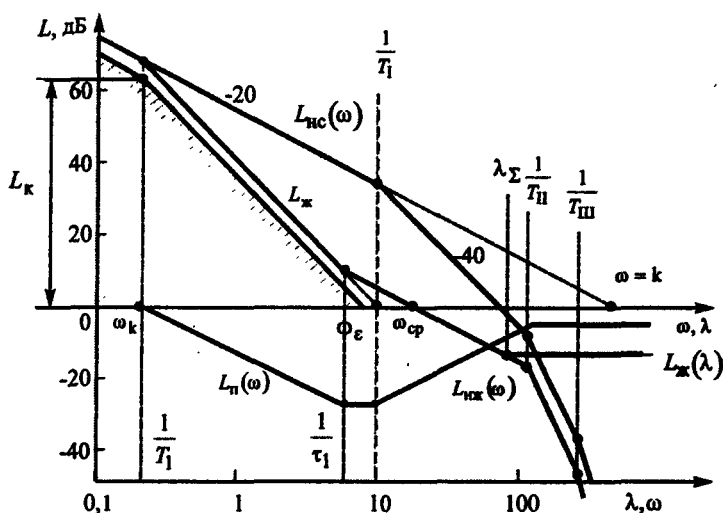


Рис. 8.4. К синтезу логарифмических частотных характеристик непрерывного корректирующего устройства

амплитудную характеристику последовательной корректирующей цепи $L_n(\omega)$, которая соответствует интегродифференцирующему звену с передаточной функцией

$$W_n(s) = \frac{(1 + \tau_{1k}s)(1 + \tau_{2k}s)}{(1 + T_{1k}s)(1 + T_{2k}s)},$$

где $T_{1k} = T_1 = 4$ с; $\tau_{1k} = \tau_1 = 0,18$ с; $T_{2k} = T_2 = 0,01$ с; $\tau_{2k} = \tau_2 = 0,1$ с.

Отметим, что при $T_\Sigma > \frac{1}{\omega_{cp}} \frac{M}{M+1}$ период дискретности, вычисленный из условия

(8.10), получится отрицательным. Это соответствует случаю, когда малые постоянные времени нескорректированной непрерывной части столь велики, что заданная точность слежения в установившемся режиме и требуемое значение показателя колебательности несовместимы. Для уменьшения малых постоянных времени непрерывной части следует ввести дополнительные корректирующие цепи, после чего, подсчитав новую величину T_Σ , определить частоту сопряжения λ_Σ желаемой ЛАЧХ импульсной системы.

Синтез параллельного корректирующего устройства W_k и корректирующей обратной связи W_∞ (см. рис. 8.1) удобно проводить, используя характеристики, или передаточную функцию эквивалентного последовательного корректирующего устройства подобно тому, как это делается для непрерывных систем, в частности:

$$W_k(s) = W_1(s)[W_n(s) - 1]; \quad W_\infty(s) = [W_n^{-1}(s) - 1]W_1^{-1}(s).$$

Приведенное выражение для передаточной функции $W_\infty(s)$ соответствует внутреннему контуру с отрицательной обратной связью.

Вариант 2. В том случае, когда период дискретности T задан, параметры желаемой характеристики в области низких частот, как и в предыдущем случае, выбираются из условий обеспечения необходимой точности в установившемся режиме и заданного значения показателя колебательности; значения k_v , постоянных времени T_1 и τ_1 , а также частоты среза системы совпадают с полученными в варианте 1.

Желаемая логарифмическая амплитудная характеристика непрерывной части в области высоких частот может быть произвольной, но такой, чтобы сумма малых постоянных времени не превышала допустимого значения:

$$T_\Sigma \leq \frac{1}{\omega_{cp}} \frac{M}{M+1} - \frac{T}{2} = 0,04 - 0,015 = 0,025 \text{ с.} \quad (8.13)$$

Для того чтобы в области высоких частот вид логарифмической характеристики $L_n(\omega)$ был близок характеристике $L_\infty(\omega)$ исходной непрерывной части, примем (см. рис. 8.4) $T_\Sigma = T_2 + T'_2 + T_3$, положив $T_2 = T_{II} = 0,01$ с, $T_3 = T_{III} = 0,005$ с. Тогда допустимое значение постоянной времени T'_2 желаемой передаточной функции непрерывной части, определяемое условием (8.13), будет:

$$T'_2 \leq 0,025 - (T_2 + T_3) = 0,01 \text{ с.}$$

Приняв $T'_2 = T_2 = T_{II} = 0,01$ с и определив значение $\varepsilon = \frac{T_\Sigma}{T} = \frac{0,025}{0,03} = 0,83$, находим последнюю частоту сопряжения желаемой ЛАЧХ импульсной системы

$$\lambda_\Sigma = \frac{2}{|1 - 2\varepsilon|T} = \frac{2}{0,63 \cdot 0,03} \approx 100 \text{ с}^{-1}.$$

По частотным характеристикам $L_{\text{н}}(\omega)$ и $L_{\text{нс}}(\omega)$ непрерывной части определим ЛАЧХ последовательной корректирующей цепи с передаточной функцией

$$W_{\text{н}}(s) = \frac{(1 + 0,18s)(1 + 0,1s)}{(1 + 4s)(1 + 0,01s)}.$$

Если сумма малых постоянных времени нескорректированной непрерывной части не соответствует условию (8.13), то необходимо либо уменьшить период дискретности, либо ввести дополнительные корректирующие цепи для уменьшения этих постоянных времени.

8.4. СИНТЕЗ ДИСКРЕТНЫХ КОРРЕКТИРУЮЩИХ УСТРОЙСТВ

Расчет дискретных корректирующих устройств в импульсных системах (см. рис. 8.2) так же, как непрерывных, удобно проводить, используя логарифмические частотные характеристики. В этом случае методика расчета дискретного последовательного корректирующего устройства полностью совпадает с методикой расчета последовательного корректирующего устройства в непрерывной системе. Логарифмическая частотная характеристика $L_{\text{н}}(\lambda)$ последовательного дискретного корректирующего устройства получается как разность между желаемой логарифмической частотной характеристикой $L_{\text{н}}(\lambda)$ и логарифмической характеристикой $L_{\text{нс}}(\lambda)$ нескорректированной системы.

По характеристике $L_{\text{н}}(\lambda)$ легко записывается частотная передаточная функция $D_{\text{н}}(j\lambda)$. Перейдя к z -преобразованию путем подстановки

$$j\lambda = \frac{2}{T} \frac{z-1}{z+1}, \quad (8.14)$$

получаем импульсную передаточную функцию $D_{\text{н}}(z)$.

При расчете дискретных корректирующих устройств необходимо учитывать некоторые особенности, которые связаны с ограничениями, накладываемыми на вид желаемой передаточной функции системы для выполнения условий реализуемости дискретной передаточной функции корректирующего устройства и грубости скорректированной системы. Рассмотрим эти особенности применительно к системе (см. рис. 8.2, а) с последовательным дискретным корректирующим устройством, передаточная функция которой в замкнутом состоянии

$$\Phi_{\text{ск}}(z) = \frac{D_{\text{н}}(z)W_{\text{нс}}(z)}{1 + D_{\text{н}}(z)W_{\text{нс}}(z)}.$$

Задача коррекции формально сводится к определению передаточной функции $D_{\text{н}}(z)$ или частотной характеристики $L_{\text{н}}(\lambda)$ корректирующего

щего устройства, обеспечивающих тождество передаточных функций замкнутых

$$\Phi_{\text{ск}}(z) \equiv \Phi_{\text{ж}}(z), \quad (8.15)$$

а значит, и разомкнутых

$$W_{\text{ск}}(z) \equiv W_{\text{ж}}(z) \quad (8.16)$$

систем. Это значит, что синтезируемая передаточная функция

$$D_{\text{п}}(z) = \frac{1}{W_{\text{нс}}(z)} \frac{\Phi_{\text{ж}}(z)}{1 - \Phi_{\text{ж}}(z)}, \quad (8.17)$$

и, следовательно,

$$D_{\text{п}}(z) = \frac{W_{\text{ж}}(z)}{W_{\text{нс}}(z)}. \quad (8.18)$$

Приведенные выражения позволяют получить условия, при которых дискретная передаточная функция $D_{\text{п}}(z)$ будет физически реализуемой (осуществимой), т.е. порядок полинома ее числителя будет не выше порядка полинома знаменателя. Действительно, если подставить в (8.17) выражение для $W_{\text{нс}}(z)$ в виде $W_{\text{нс}}(z) = \frac{\Phi(z)}{1 - \Phi(z)}$, где $\Phi(z)$ — передаточная функция исходной замкнутой системы, то получим

$$D_{\text{п}}(z) = \frac{1 - \Phi(z)}{\Phi(z)} \frac{\Phi_{\text{ж}}(z)}{1 - \Phi_{\text{ж}}(z)}.$$

На основе этого выражения легко сделать вывод о том, что для осуществимости $D_{\text{п}}(z)$ необходимо, чтобы разность степеней полиномов знаменателя и числителя желаемой передаточной функции $\Phi_{\text{ж}}(z)$ была не меньше соответствующей разности степеней полиномов $\Phi(z)$.

Соответствующее требование к передаточной функции $W_{\text{ж}}(z)$ можно получить из выражения (8.18), если обозначить

$$W_{\text{нс}}(z) = \frac{B(z)}{A(z)}; \quad W_{\text{ж}}(z) = \frac{B_{\text{ж}}(z)}{A_{\text{ж}}(z)}. \quad (8.19)$$

Подставив (8.19) в (8.18), запишем:

$$D_{\text{п}}(z) = \frac{B_{\text{ж}}(z)A(z)}{B(z)A_{\text{ж}}(z)},$$

откуда получаем искомое условие физической реализуемости $D_{\text{п}}(z)$:

$$m_{\text{ж}} + n \leq m + n_{\text{ж}},$$

где $m_{\text{ж}}$, m — степени полиномов $B_{\text{ж}}(z)$, $B(z)$; $n_{\text{ж}}$, n — степени полиномов $A_{\text{ж}}(z)$, $A(z)$, или

$$n_x - m_x \geq n - m. \quad (8.20)$$

Очевидно, что такое же требование должно выполняться и при синтезе последовательного корректирующего устройства в непрерывных системах. Если обратить внимание на то, что высокочастотная асимптота ЛАЧХ имеет наклон $-20(n-m)$ дБ/дек, то из (8.20) следует требование, накладываемое на желаемую частотную характеристику $L_x(\lambda)$: отрицательный наклон высокочастотной асимптоты этой ЛАЧХ должен быть не меньше наклона соответствующей асимптоты $L_{ис}(\lambda)$ ЛАЧХ исходной системы.

Выполнение условий осуществимости передаточной функции $D_n(z)$, обеспечивающей формально тождества (8.15), (8.16), не является единственным требованием для достижения заданных показателей качества замкнутой системы. Или, другими словами, даже если желаемая частотная характеристика $L_x(\lambda)$ либо желаемая передаточная функция $W_x(z)$ выбраны исходя из этих показателей качества, а также реализованы тождества (8.15) или (8.16) и передаточная функция $D_n(z)$ осуществима, может оказаться, что замкнутая система будет негрубой и неработоспособной. Такая особенность имеет место, если неизменяемая часть системы с передаточной функцией $W_{ис}(z)$ (исходная система в разомкнутом состоянии) не является минимально-фазовой, т.е. содержит нули или полюсы по модулю большие единицы (вне единичного круга). Поясним это, по-прежнему имея в виду дискретную коррекцию последовательного типа.

Предположим, что желаемая ЛАЧХ $L_x(\lambda)$ построена в соответствии с методикой, изложенной в § 8.2 и записана соответствующая передаточная функция $W_x(z)$. Деформация ЛАЧХ исходной системы $L_{ис}(\lambda)$ для достижения тождеств $L_{ск}(\lambda) \equiv L_x(\lambda)$; $W_{ск}(z) \equiv W_x(z)$ связана, в частности, с изменением наклонов асимптот $L_{ис}(\lambda)$ за некоторыми ее частотами сопряжения, т.е. компенсацией каких-либо нулей или полюсов передаточной функции $W_{ис}(z)$ равными им соответствующими полюсами или нулями $D_n(z)$.

Обозначим часть передаточной функции $W_{ис}(z)$, содержащей указанные компенсируемые нули или полюсы, как $W_k(z) = B_1(z)/A_1(z)$, т.е. запишем:

$$W_{ис}(z) = W_{ис}^0(z) W_k(z) = W_{ис}^0(z) \frac{B_1(z)}{A_1(z)}.$$

Передаточная функция скорректированной системы в разомкнутом состоянии будет иметь вид

$$\begin{aligned} W_{ск}(z) &= D_n^0(z) \frac{A_1(z)}{B_1(z)} W_{ис}^0(z) \frac{B_1(z)}{A_1(z)} = \\ &= D_n^0(z) W_{ис}^0(z) \frac{A_1(z) B_1(z)}{B_1(z) A_1(z)}, \end{aligned} \quad (8.21)$$

где $D_n^0(z)W_{nc}^0(z) \equiv W_{\Sigma}(z)$. Таким образом, передаточная функция $W_{\Sigma}(z)$ тождественна желаемой $W_{\Sigma}(z)$ с точностью до «неполной части» скорректированного контура, имеющей характеристический полином (ХП)

$$A_n(z) = B_1(z)A_1(z).$$

При замыкании системы с передаточной функцией (8.21) единичной отрицательной обратной связью ее характеристическое уравнение принимает вид

$$1 + W_{\Sigma}(z) \frac{A_1(z)B_1(z)}{B_1(z)A_1(z)} = 0,$$

т.е. ХП скорректированной замкнутой системы

$$A_2(z) = [A_{\Sigma}(z) + B_{\Sigma}(z)]A_1(z)B_1(z) = A_{\Sigma}(z)A_n(z). \quad (8.22)$$

Как видно, ХП замкнутой системы представляется в виде произведения двух полиномов: $A_{\Sigma}(z)$ — желаемый ХП, расположение корней которого обеспечивает заданные динамические свойства системы управления; $A_n(z)$ — ХП «неполной части» скорректированного контура. «Неполная часть» содержит те нули и полюсы передаточной функции $W_{nc}(z)$, которые не повторяются в желаемой передаточной функции $W_{\Sigma}(z)$, т.е. в нее не входят. Поскольку, как это видно из (8.22), все нули и полюсы $W_{nc}(z)$, которые не вошли в $W_{\Sigma}(z)$, после замыкания системы становятся корнями ее ХП, поэтому они при замыкании не перемещаются.

Отсюда следует принципиальный вывод — все нули и полюсы «вне круга» передаточной функции $W_{nc}(z)$ («плохие» ее нули и полюсы) должны обязательно входить в $W_{\Sigma}(z)$, т.е. их нельзя компенсировать соответствующими полюсами и нулями $D_n(z)$.

Рассмотрим определение передаточной функции дискретного корректирующего устройства для импульсной следящей системы из примера, приведенного в § 8.3, для случая, когда период дискретности задан.

Дискретная частотная передаточная функция исходной системы $W_{nc}(j\lambda)$ с непрерывной частью вида (8.11) и соответствующей логарифмической амплитудно-частотной характеристикой $L_{nc}(\omega)$ (см. рис. 8.4), имеющей наклон -40 дБ/дек асимптоты в окрестности ω_{cp} , имеет вид

$$W_{nc}(j\lambda) = W_n(s) \Big|_{s=j\lambda} \frac{1-j\lambda \frac{T}{2}}{1+j\lambda \frac{T}{2}} \left[1 + j\lambda \left(\frac{T}{2} - T_{\Sigma} \right) - (j\lambda)^2 \left(\frac{TT_{\Sigma}}{2} - T_{\sigma}^2 \right) \right], \quad (8.23)$$

где $W_n(s) = \frac{k}{s(T_1s+1)}$; $T_{\Sigma} = T_{II} + T_{III} = 0,015$ с; $T_{\sigma}^2 = T_{II}^2 + T_{III}^2 + T_{II}T_{III} = 0,000175$ с².

Допустим, что задано значение периода дискретности системы $T = 0,02$ с. Нетрудно убедиться в том, что при исходных параметрах системы выполняется условие $T_0^2 \ll TT_\Sigma/2$, что позволяет упростить выражение (8.23), переписав его в виде

$$W_{\text{ис}}(j\lambda) \cong \frac{k \left(1 - j\lambda \frac{T}{2}\right) (1 - j\lambda T_\Sigma)}{j\lambda (1 + j\lambda T_I)} = \frac{k(1 - j\lambda 0,01)(1 - j\lambda 0,015)}{j\lambda (1 + j\lambda 0,1)}.$$

Поскольку частотная передаточная функция исходной системы имеет два правых нуля, что соответствует двум правым нулям в импульсной передаточной функции

$$W_{\text{ис}}(w) = W_{\text{ис}}(j\lambda) \Big|_{j\lambda = \frac{2}{T}w}$$

и двум нулям передаточной функции

$$W_{\text{ис}}(z) = W_{\text{ис}}(w) \Big|_{w = \frac{z-1}{z+1}},$$

лежащим вне круга единичного радиуса на плоскости z , необходимо так сформировать желаемую передаточную функцию $W_{\text{ж}}(j\lambda)$, чтобы она также включала в себя эти нули.

В области низких частот ($\lambda < \lambda_{\text{cp}}$) сохраним вид и параметры желаемой характеристики, принятые в примере из § 8.3. Введя дополнительный полюс для выполнения условия физической реализуемости передаточной функции $W_{\text{ж}}$, получим

$$W_{\text{ж}}(j\lambda) = k \frac{(1 + j\lambda \tau_1) \left(1 - j\lambda \frac{T}{2}\right) (1 - j\lambda T_\Sigma)}{j\lambda (1 + j\lambda T_I) (1 + j\lambda T_\Sigma)}.$$

Для обеспечения требуемого значения показателя колебательности M в этом случае необходимо выполнить условие

$$\frac{T}{2} + 2T_\Sigma \leq \frac{1}{\lambda_{\text{cp}}} \frac{M}{M+1}, \quad \text{или} \quad T_\Sigma \leq \frac{1}{2\lambda_{\text{cp}}} \frac{M}{M+1} - \frac{T}{4}.$$

Подставив значения $T = 0,02$ с; $M = 1,5$; $\lambda_{\text{cp}} = 15,3$ с⁻¹, убеждаемся, что принятое значение $T_\Sigma = T_{\text{II}} + T_{\text{III}}$ этому условию удовлетворяет.

На рис. 8.5 построены желаемая ЛАЧХ $L_{\text{ж}}(\lambda)$ и ЛАЧХ исходной системы $L_{\text{ис}}(\lambda)$. По виду разностной ЛАЧХ $L_{\text{н}}(\lambda)$ легко записать частотную передаточную функцию последовательного корректирующего устройства:

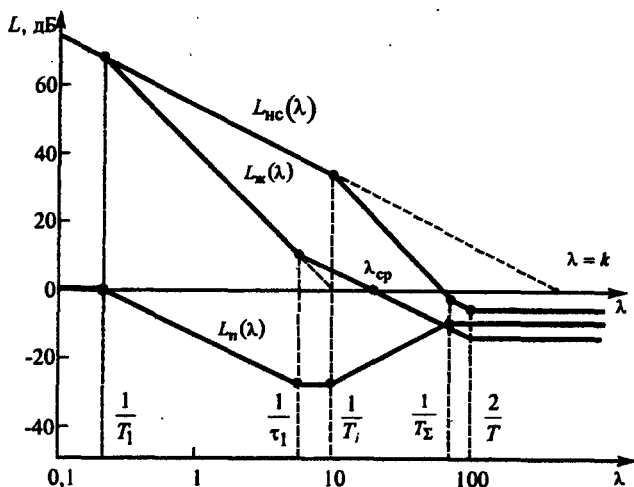


Рис 8.5. К синтезу логарифмических частотных характеристик дискретного корректирующего устройства

$$D_n(j\lambda) = \frac{W_k(j\lambda)}{W_{nc}(j\lambda)} = \frac{(1 + j\lambda\tau_1)(1 + j\lambda T_l)}{(1 + j\lambda T_1)(1 + j\lambda T_2)}$$

Используя подстановку (8.14), получим

$$D_n(z) = \frac{b_2 z^2 + b_1 z + b_0}{a_2 z^2 + a_1 z + a_0},$$

где

$$b_0 = \left(1 - 2\frac{\tau_1}{T}\right) \left(1 - 2\frac{T_l}{T}\right); \quad b_1 = \left(1 + 2\frac{\tau_1}{T}\right) \left(1 - 2\frac{\tau_1}{T}\right) + \left(1 - 2\frac{\tau_1}{T}\right) \left(1 + 2\frac{T_l}{T}\right);$$

$$b_2 = \left(1 + 2\frac{\tau_1}{T}\right) \left(1 + 2\frac{T_l}{T}\right); \quad a_0 = \left(1 - 2\frac{T_l}{T}\right) \left(1 - 2\frac{T_2}{T}\right);$$

$$a_1 = \left(1 + 2\frac{T_l}{T}\right) \left(1 - 2\frac{T_2}{T}\right) + \left(1 + 2\frac{T_2}{T}\right) \left(1 - 2\frac{T_l}{T}\right); \quad a_2 = \left(1 + 2\frac{T_l}{T}\right) \left(1 + 2\frac{T_2}{T}\right).$$

Корректирующее устройство с передаточной функцией $D_n(z)$ может быть реализовано либо в виде импульсной RC-цепи, либо — цифрового вычислителя. В последнем случае управляющая последовательность $\mu(n)$; $n = 0, 1, 2, \dots$ на его выходе должна вычисляться по рекуррентной формуле

$$\mu(n) = \frac{1}{a_2} [b_2 e(n) + b_1 e(n-1) + b_0 e(n-2) - a_1 \mu(n-1) - a_0 \mu(n-2)].$$

Глава 9

НЕЛИНЕЙНЫЕ МОДЕЛИ СИСТЕМ УПРАВЛЕНИЯ

9.1. НЕОБХОДИМОСТЬ В НЕЛИНЕЙНЫХ МОДЕЛЯХ

Расчет систем управления по нелинейным моделям значительно сложнее, чем по линейным. Это объясняется большим разнообразием движений, описываемых нелинейными уравнениями. Переход от линейных моделей к нелинейным, т. е. их усложнение — мера вынужденная; необходимость расширения и углубления знаний о поведении систем управления должна быть обоснована. Термину «нелинейный» обычно придают расширенный смысл: «не обязательно линейный».

Напомним, что решение обыкновенных линейных дифференциальных уравнений с постоянными детерминированными коэффициентами

$$d_n \frac{d^n x}{dt^n} + \dots + d_1 \frac{dx}{dt} + d_0 x = b_m \frac{d^m f}{dt^m} + \dots + b_1 \frac{df}{dt} + b_0 f \quad (9.1)$$

представляется в виде суммы

$$x(t) = x_{\text{св}}(t) + x_{\text{св}}(t),$$

где $x_{\text{св}}(t)$ — частное решение неоднородного уравнения (9.1), описывающее вынужденное движение;

$$x_{\text{св}}(t) = \sum_{i=1}^n C_i e^{p_i t} \quad (9.2)$$

— общее решение соответствующего однородного уравнения, описывающее свободное движение при начальных условиях: $x(0)$, $x'(0), \dots, x^{(n-1)}(0)$. Выражение (9.2) записано для случая простых корней $p_1, \dots, p_n$ характеристического полинома

$$D(p) = d_n p^n + \dots + d_1 p + d_0. \quad (9.3)$$

Относительная простота анализа линейных моделей объясняется возможностью отдельного анализа вынужденных и свободных движений, а также тем, что известна форма решения (9.2). Построение решения сводится к алгебраическим задачам вычисления корней харак-

теристического полинома и решения системы линейных уравнений относительно коэффициентов C_i .

Вместе с тем, разнообразие движений, описываемых линейными уравнениями, может оказаться недостаточным. Повышая порядок n уравнений и подбирая коэффициенты, не всегда удастся объяснить реальные процессы на больших интервалах времени и в широких диапазонах амплитуд переменных. Рассмотрим некоторые из таких примеров.

Неединственность положения равновесия. Линейная система имеет единственное положение равновесия. Например, выходная переменная системы, описываемой уравнением (9.1), при постоянном воздействии f принимает единственное значение, равное

$$x = \frac{b_0}{d_0} f.$$

Если система находится не на границе устойчивости, то при ненулевых начальных условиях движения асимптотически затухают к положению равновесия (система устойчива в целом) или расходятся (система не устойчива).

Реальные динамические системы могут иметь несколько положений равновесия. Одним из примеров является маятник (см. рис. 1.7, а). Положения равновесия маятника образуют счетное множество $\{y = k\pi; k = 0, \pm 1, \pm 2, \dots\}$, причем, четные k отвечают верхним, а нечетные — нижним положениям равновесия.

В качестве другого примера представим себе следящую систему управления, датчик рассогласования и электродвигатель привода которой имеют зоны нечувствительности. Если рассогласование мало, то напряжение на входе двигателя по модулю меньше напряжения трогания и двигатель не вращается. Положения равновесия системы относительно переменной входа двигателя образуют отрезок, иначе множество мощностей континуума.

Конечная длительность процессов. Если линейная система устойчива, т. е. корни p , характеристического полинома (9.3) имеют отрицательные действительные части, то соответствующие экспоненты в решении (9.2) затухают в бесконечности. Реально же длительность процессов управления конечна (пунктирная линия на рис. 9.1), чему способствуют зоны нечувствительности элементов, сухое трение и люфты (зазоры) кинематических сочленений.

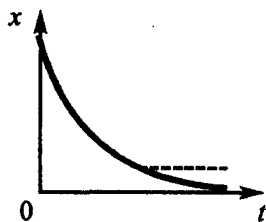


Рис.9.1. Конечная длительность процессов

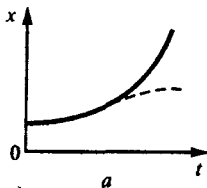
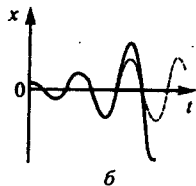


Рис.9.2. Ограниченность уровней переменных



Ограниченность уровней переменных. Если линейная система неустойчива, то значения переменных неограниченно растут. Реально уровни переменных всегда ограничены энергетическими, материальными, прочностными ресурсами. На рис. 9.2 сплошными линиями показаны экспоненциально расходящиеся процессы в линейных системах первого (рис. 9.2, а) и второго (рис. 9.2, б) порядков. Пунктирные линии соответствуют реальным процессам.

Автоколебания — периодические движения за счет внутренних свойств системы при отсутствии внешних колебательных воздействий. В линейных системах периодические движения гармонической формы соответствуют колебательной границе устойчивости. Амплитуды этих колебаний зависят от начальных условий. При самом незначительном изменении параметров системы колебания превращаются в затухающие или расходящиеся.

Автоколебания имеют относительно стабильные амплитуду и частоту, которые восстанавливаются после снятия возмущений. На рис. 9.2, б пунктирная линия соответствует автоколебаниям в системе с неустойчивым положением равновесия.

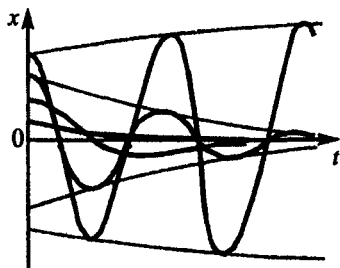


Рис.9.3. Зависимость процессов от начальных условий

Зависимость характера движений от начальных условий и уровней воздействий. В реальных системах не выполняется принцип суперпозиции — при сложении воздействий реакция не равна сумме реакций на отдельные воздействия. На рис. 9.3 показаны графики процессов в одной и той же динамической системе (физический маятник) в зависимости от начальных условий.

Примеры явлений, не объясняемых теорией линейных систем, можно продолжить [1, 15, 46, 52, 58, 72, 75, 78].

Нелинейные математические модели, используемые для анализа систем управления, появляются вследствие учета естественных (сопутствующих) эффектов, присущих объекту или элементам системы управления и обусловленных нелинейным характером законов природы, которым подчиняются исследуемые явления. Нелинейности могут вводиться и специально с целью компенсации нежелательных эффектов от естественных нелинейностей или для придания системе управления особых свойств, которые принципиально недостижимы линейными средствами. Так, именно нелинейные алгоритмы управления могут обеспечить максимальное быстроедействие процессов при наличии естественных ограничений на уровни управляющих воздействий; нелинейности обязательно вводятся при создании генераторов колебаний и т. д. В ряде систем управления техническими объектами нелинейные, в частности, релейные регулирующие устройства оказываются наиболее простыми, дешевыми и надежными.

9.2. БЕЗЫНЕРЦИОННЫЕ НЕЛИНЕЙНЫЕ ЭЛЕМЕНТЫ

В теории и практике управления элементы и системы рассматривают как преобразователи сигналов — носителей информации о цели, состоянии объекта и воздействиях среды (рис. 9.4).

Как известно, линейный безынерционный элемент полностью задается значением его коэффициента усиления. Нелинейные зависимости между постоянными значениями входных и выходных сигналов

$$y = F(x)$$

могут задаваться аналитически, графически или таблично. В том случае, когда нелинейный элемент (НЭ) имеет один вход и один выход особенно наглядны графики статических характеристик (СХ) (рис. 9.5).

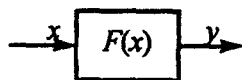


Рис. 9.4. Элемент системы как преобразователь сигнала

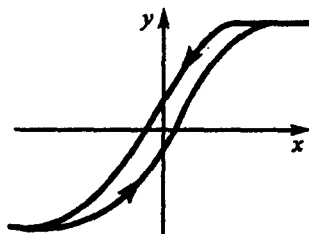


Рис. 9.5. Пример статической характеристики НЭ

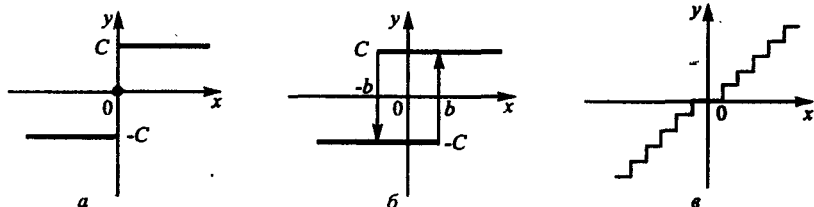


Рис. 9.6. Кусочно-постоянные (релейные) СХ

Условия преобразования сигналов безынерционными НЭ зависят от уровней сигналов и не зависят от их частоты.

Приведем некоторые примеры безынерционных НЭ и их СХ.

Пример 1. *Нелинейные элементы с кусочно-постоянными СХ.* Простейшим представителем нелинейностей этой группы является так называемое идеальное реле (рис. 9.6, а):

$$y = C \operatorname{sgn} x = \begin{cases} C, & x > 0; \\ 0, & x = 0; \\ -C, & x < 0. \end{cases}$$

Более тонкое изучение может показать, что релейное устройство имеет гистерезис (рис. 9.7, б). Выражение для двузначной СХ с разрывами первого рода можно записать так:

$$y = \begin{cases} C, & x > b; \\ -C, & x < -b; \\ C, & |x| \leq b, y_0 = C; \\ -C, & |x| \leq b, y_0 = -C, \end{cases}$$

где b — половина зоны неоднозначности СХ; y_0 — состояние реле, равное значению y до входа в зону неоднозначности. Таким образом, этот безынерционный НЭ обладает памятью — значение его выхода определяется не только значением входа в тот же момент, но также и предысторией (состоянием) НЭ по уровню сигнала.

Другим примером НЭ с кусочно-постоянной однозначной СХ является квантование сигналов по уровню в преобразователях аналог-код, предназначенных для ввода информации о состоянии непрерывных процессов в цифровые управляющие устройства (рис. 9.6, в). Малая разрядность ЭВМ может оказаться существенным препятствием к достижению высокой точности и хорошего качества процессов в окрестности положений равновесия.

Пример 2. *Нелинейные элементы с кусочно-линейными СХ.* На рис. 9.7, а показан график СХ НЭ типа «насыщение»:

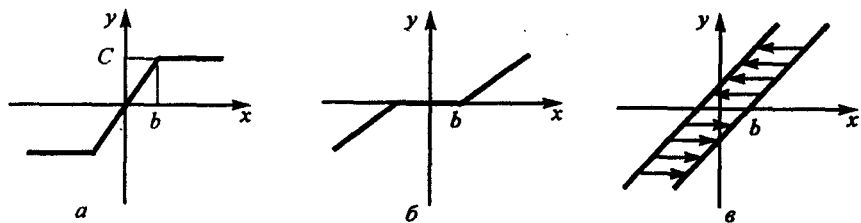


Рис. 9.7. Кусочно-линейные СХ

$$y = \begin{cases} C, & x \geq b; \\ -C, & x \leq -b; \\ (C/b)x, & |x| < b. \end{cases}$$

Как правило, эта нелинейность вводится в модели для учета ограничений уровней переменных при исследовании поведения систем управления в режимах больших отклонений от положения равновесия.

Нелинейный элемент типа «зона нечувствительности» (рис. 9.7, б) учитывает реальные свойства датчиков, исполнительных механизмов и других устройств при малых входных сигналах.

Нелинейность типа «люфт» (рис. 9.7, в) является многозначной — одному значению входа соответствует бесчисленное множество (континуум) значений выхода. Этот НЭ моделирует кинематические сочленения механических приборов и устройств (например, редукторов).

Приведенные выше примеры кусочно-линейных СХ непрерывны, но имеют разрыв производной dy/dx . Могут быть и кусочно-линейные СХ с разрывами первого рода.

Пример 3. Нелинейные элементы с гладкими СХ. Гладкие СХ имеют непрерывные производные. Примерами являются характеристики термопары (рис. 9.8, а), устройства возведения входного сигнала в квадрат (рис. 9.8, б), в куб (рис. 9.8, в), индукционных электромеханических преобразователей угла, электромагнитных явлений с гистерезисом и др.

Нелинейные зависимости между значениями входа и выхода одномерных НЭ можно задавать параметрически — парой функций $\langle (x(t), y(t)) \rangle$;

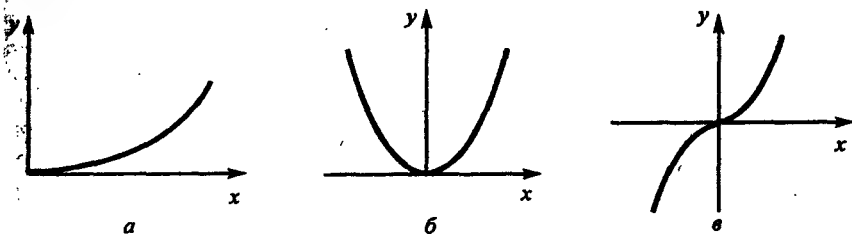


Рис. 9.8. Гладкие СХ

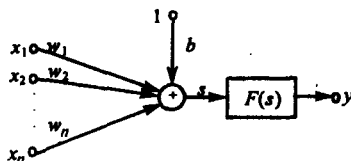


Рис. 9.9. Модель нейрона

исключая параметр t , получим непосредственную связь между переменными входа и выхода. В случае однозначных СХ в качестве входа $x(t)$ особенно удобен периодический сигнал треугольной формы с достаточной амплитудой — выход НЭ будет периодически повторять форму СХ. Для сложных НЭ с неоднозначными СХ выбор функции $x(t)$ из условия исчерпывающего задания НЭ парой вход-выход является нетривиальной задачей. По существу, речь идет об экспериментальном исследовании НЭ, успех которого зависит от априорной информации.

Пример 4. *Нейронные сети как многомерные нелинейные элементы.* Элементы (объекты, системы) называют *многомерными*, если переменные входа и/или выхода являются векторами. Линейные безынерционные многомерные элементы задаются матрицами коэффициентов усиления. В нелинейном случае необходимы другие способы описания.

В последние годы для аппроксимации нелинейных зависимостей часто применяют искусственные нейронные сети — многомерные настраиваемые безынерционные преобразователи [74, 94, 97].

На рис. 9.9 изображена модель нейрона. Выходная переменная y является нелинейной функцией взвешенной суммы входных переменных

$$y = F \left(b + \sum_{i=1}^n w_i x_i \right)$$

Параметры нейрона — смещение b и весовые коэффициенты w_i можно настраивать, добиваясь требуемой зависимости выхода от входов.

Широкие аппроксимирующие возможности достигаются, если нейроны образуют сети определенной архитектуры. На рис. 9.10 изображена однонаправленная двухслойная нейронная сеть с тремя входами, двумя выходами и тремя нейронами в первом (скрытом) слое. Весовые коэффициенты и смещения первого слоя упорядочены в матрицы W_1, b_1 , а второго — в матрицы W_2, b_2 .

Разработан ряд алгоритмов обучения многослойных сетей — настройки весов и смещений из условия минимизации функционала, оценивающего сумму квадратов отклонений выхода сети от выхода моделируемого объекта.

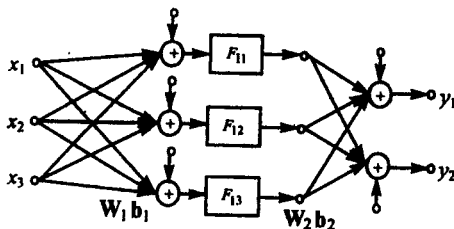


Рис. 9.10. Пример двухслойной нейронной сети

Для аппроксимации многомерных нелинейностей также применяют элементы, реализующие так называемую нечеткую логику [95, 96, 97].

9.3. ДИНАМИЧЕСКИЕ НЕЛИНЕЙНЫЕ ЭЛЕМЕНТЫ

В общем случае дифференциальные уравнения, описывающие элементы систем или сами системы, являются нелинейными

$$\Phi(y, y', \dots, y^{(n)}; x, x', \dots, x^{(m)}) = 0. \quad (9.4)$$

Иногда они разрешаются относительно старшей производной переменной выхода

$$y^{(n)} = \varphi(y, y', \dots, y^{(n-1)}; x, x', \dots, x^{(m)}). \quad (9.5)$$

Примерами служат дифференциальные уравнения математического маятника

$$y'' + \omega_0^2 \sin y = 0$$

и уравнение Ван дер Поля

$$y'' - \mu(1 - y^2)y' + y = 0.$$

Часто дифференциальные уравнения представляются в форме Коши:

$$\begin{aligned} \frac{dv}{dt} &= \varphi(v, x); \\ y &= \psi(v, x), \end{aligned} \quad (9.6)$$

где v — вектор переменных состояния; φ — вектор-функция; ψ — функция выхода. В уравнениях (9.4) — (9.6) предполагается, что нелинейные функции заданы аналитически.

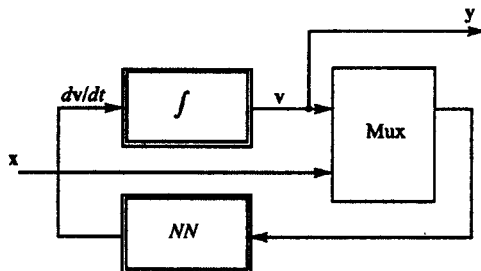


Рис. 9.11. Реализация нелинейной динамической модели

Временная характеристика динамического линейного элемента — функция веса $w(t)$ — позволяет связывать переменные входа и выхода с помощью интеграла свертки (2.5). В ряде случаев нелинейные преобразования описываются функционалами типа Вольтерра (V. Volterra), являющимися обобщением интеграла свертки. Представление НЭ функционалами Вольтерра позволяет распространить понятие передаточной функции на нелинейные преобразования, а также развить спектральный метод анализа нелинейных систем [75].

В линейных динамических элементах условия преобразования сигналов определялись лишь частотным спектром сигнала и не зависели от его уровня. Преобразование сигналов динамическими НЭ в значительной степени зависит как от уровней сигналов, так и от их частотных спектров.

Для описания динамических нелинейных преобразований нейронные сети дополняются интеграторами в случае непрерывных систем и элементами сдвига — в случае дискретных систем.

На рис. 9.11 изображена обобщенная структура, позволяющая реализовать нелинейную модель в форме Коши.

На рис. 9.11 блок NN означает многомерную нейронную сеть, а блок Mux — объединяет векторы состояния v и входа x .

9.4. НЕЛИНЕЙНЫЕ МОДЕЛИ С РАСКРЫТОЙ СТРУКТУРОЙ

Во многих случаях нелинейные модели появляются в результате дополнения линейных моделей нелинейными элементами, учитывающими такие естественные факторы, как ограниченность управляющих воздействий, наличие зоны нечувствительности в измерительных и исполнительных элементах, люфтов в кинематических сочленениях или искусственное введение нелинейностей в алгоритмы управления для

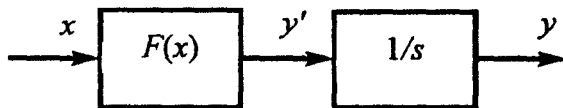


Рис. 9.12. Нелинейный интегратор

получения свойств, не достижимых в линейных системах. Простейший пример такой модели — нелинейный интегратор

$$\frac{dy}{dt} = F(x)$$

структурно изображается как последовательное соединение безынерционного НЭ и линейного интегрирующего звена (рис. 9.12). На рис. 9.13, а изображен другой пример — модель системы с обратной связью в форме структурной схемы. На рис. 9.13, б та же модель представлена в форме сигнального графа, одна из дуг которой помечена двумя штрихами, указывающими на нелинейный характер преобразования сигнала.

В этих примерах разделены динамическая линейная часть (ЛЧ) и безынерционная нелинейность — нелинейные эффекты сосредоточены в безынерционном, а динамические — в линейном элементах. Такое же разделение иллюстрирует схема реализации многомерного динамического нелинейного объекта с помощью нейронной сети (см. рис. 9.11).

9.5. РАСЧЕТНЫЕ ФОРМЫ НЕЛИНЕЙНЫХ МОДЕЛЕЙ

Класс нелинейных моделей очень широк, что чрезвычайно затрудняет их единообразное описание, возможность использования универсальных методов анализа и синтеза. Поэтому при разработке методик исследования автоматических систем управления по нелинейным моделям выбираются некоторые расчетные формы моделей, к которым, по возможности, пытаются привести исходные. Основными соображениями при выборе расчетных форм нелинейных моделей являются

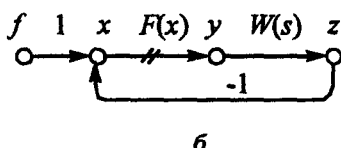
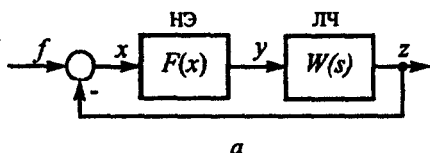


Рис. 9.13. Нелинейная система с обратной связью.
Расчетная структура нелинейных моделей

широта класса автоматических систем, моделируемых принятыми типами моделей, удобство последующего анализа и простота эквивалентного приведения к этому виду других нетиповых исходных моделей. Ниже рассматриваются некоторые наиболее часто используемые формы представления нелинейных моделей расчетных видов и обсуждаются вопросы взаимных переходов между этими формами.

Большие достижения имеются в области теории нелинейных систем, описываемых дифференциальными уравнениями второго порядка:

$$x'' = \varphi(x, x')$$

или в форме Коши:

$$\begin{aligned} dv_1 / dt &= \varphi_1(v_1, v_2); \\ dv_2 / dt &= \varphi_2(v_1, v_2), \end{aligned} \quad (9.7)$$

где v_1, v_2 — переменные состояния (фазовые координаты) динамической системы. От дифференциального уравнения второго порядка, разрешенного относительно старшей производной, легко перейти к форме Коши, если положить $v_1 = x, v_2 = x'$

$$\begin{aligned} dx / dt &= x', \\ dx' / dt &= \varphi(x, x'). \end{aligned}$$

В правых частях приведенных уравнений не содержится время t в явном виде; такие системы называются *автономными*. Автономная система не имеет связей со средой — отсутствуют внешние воздействия, а параметры не изменяются во времени, т. е. она *стационарна*.

Следует подчеркнуть, что хотя нелинейные дифференциальные уравнения второго порядка образуют весьма малую часть возможных нелинейных моделей, они позволяют выявить многие особенности поведения автоматических систем, не поддающиеся объяснению в рамках линейной теории.

Большая группа точных и приближенных методик исследования автоматических систем на основе нелинейного подхода использует типовую структурную схему (см. рис.9.13), моделирующую широкий класс систем управления с обратной связью. В качестве дополнительного ограничения этого расчетного вида математических моделей обычно принимают безынерционность единственного НЭ с одним входом x и одним выходом y . Динамические свойства системы сосредоточены в линейной части с передаточной функцией

$$W(s) = \frac{B(s)}{D(s)} = \frac{b_0 + b_1 s + \dots + b_m s^m}{d_0 + d_1 s + \dots + d_n s^n}. \quad (9.8)$$

По структурной схеме или графу расчетного вида (см. рис. 9.13) легко записать уравнения:

$$\begin{aligned} D(p)z(t) &= B(p)y(t); \\ y(t) &= F(x(t)); \\ x(t) &= f(t) - z(t), \end{aligned} \quad (9.9)$$

где $p \equiv d/dt$ — оператор дифференцирования. Для автономной системы воздействие тождественно нулю: $f(t) \equiv 0$, а уравнения (9.9) приводятся к виду:

$$D(p)x(t) + B(p)F(x(t)) = 0.$$

Уравнения ЛЧ в форме пространства состояний при условии $m < n$ (т. е. степень полинома числителя передаточной функции ЛЧ (9.8) ниже степени полинома знаменателя) имеют вид:

$$dv/dt = Av + By;$$

$$z = Cv.$$

С учетом НЭ для автономной системы получим следующие уравнения:

$$dv/dt = Av + BF(-Cv);$$

$$z = Cv.$$

Автономная система расчетной структуры может быть представлена и в форме интегрального уравнения

$$x(t) = - \int_0^t w(\tau) F(x(t-\tau)) d\tau,$$

где $w(t)$ — функция веса ЛЧ.

Если структурная схема исследуемой системы отличается от расчетной, то в некоторых случаях ее можно преобразовать в расчетную форму. Проблема заключается в том, что в случае нелинейных моделей не всегда применимы правила структурных преобразований линейных систем.

Рассмотрим один из случаев, когда нетиповую структуру можно преобразовать в расчетную форму. Пусть модель системы имеет сколько угодно сложную ЛЧ и единственный НЭ, с одним входом и одним выходом, как это условно показано в виде диаграммы графа на рис. 9.14. Представим ЛЧ с двумя входами (y, f) и двумя выходами (x, z) биграфом (см. пунктирные линии на рис. 9.14), для чего вычислим элементы передаточной матрицы (см. § 2.11):

$$\begin{pmatrix} \Phi_{xy} & \Phi_{xf} \\ \Phi_{zy} & \Phi_{zf} \end{pmatrix}.$$

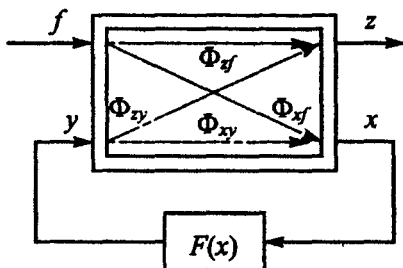
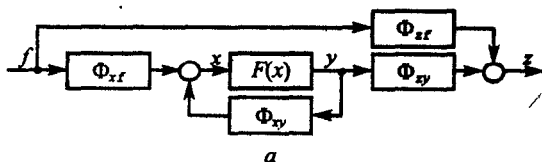
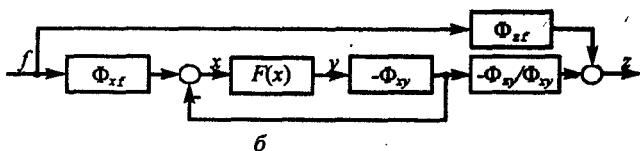


Рис. 9.14. Система с единственным НЭ и произвольной ЛЧ



а



б

Рис. 9.15. Расширенная расчетная структура

Представим диаграмму графа в виде структурной схемы, изображенной на рис. 9.15, а; после переноса точки съема получим искомую структурную схему (рис. 9.15, б) — расширенную расчетную форму, которая кроме основного контура с НЭ содержит линейные дуги, включенные последовательно с основным контуром (Φ_{xf} , $-\Phi_{xy}/\Phi_{xy}$) и параллельно с ним (Φ_{xy}). Поскольку дополнительные линейные дуги не образуют контуров, то по результатам исследования основного контура с НЭ (см. рис. 9.13) легко получить необходимые суждения о свободных и вынужденных процессах в системе со структурой, показанной на рис. 9.15.

Глава 10

АНАЛИЗ РАВНОВЕСНЫХ РЕЖИМОВ

10.1. РАВНОВЕСНЫЕ РЕЖИМЫ

Режим функционирования системы управления называется *равновесным*, если ее переменные не изменяются во времени. Для большинства систем управления промышленными и другими объектами равно-

весные режимы при постоянных внешних воздействиях являются, как правило, оптимальными в смысле принятых технологических критериев. Поэтому анализ и синтез систем по требованиям к этим режимам являются первоочередными задачами проектирования.

В равновесных режимах производные по времени равны нулю для всех переменных. Если, например, зависимость между переменными входа $f(t)$ и выхода $y(t)$ системы управления описывается дифференциальным уравнением

$$\Phi(y, y', \dots, y^{(n)}; f, f', \dots, f^{(m)}) = 0, \quad (10.1)$$

то частная модель для равновесных режимов ($f = \text{const}, y = \text{const}$) получится приравниванием нулю производных

$$\Phi(y, f) = 0.$$

Из этой неявной зависимости численным методом находят значения y при заданных значениях f (если решение существует).

Для линейных систем определение единственного положения равновесия при заданных воздействиях или определение коэффициента усиления сводится к решению систем линейных уравнений. Можно ту же задачу решать, получив передаточную функцию системы $\Phi_{\varphi}(s)$ и найдя искомый коэффициент усиления как частный случай $k_{\varphi} = \Phi_{\varphi}(0)$.

В случае нелинейных моделей, задача анализа равновесных режимов сложнее, что связано с решением систем нелинейных уравнений. Прежде всего следует ответить на вопрос о существовании решения; далее нужно определить число положений равновесия, после чего уточнить их координаты. Для решения этих задач привлекаются частные модели, по которым аналитически, графическими построениями или/и численными процедурами находятся искомые режимы и статические характеристики нелинейных систем.

10.2. ОПРЕДЕЛЕНИЕ РАВНОВЕСНЫХ РЕЖИМОВ ПО ДИФФЕРЕНЦИАЛЬНЫМ УРАВНЕНИЯМ

Выше рассмотрен случай определения положений равновесия, если имеется дифференциальное уравнение (10.1), непосредственно связывающее переменные входа и выхода системы. В общем случае объекты или системы управления описываются системами дифференциальных уравнений, например, приведенных к форме Коши:

$$\begin{aligned}\frac{dv}{dt} &= \Phi(v, f); \\ y &= \Psi(v, f).\end{aligned}\tag{10.2}$$

Система уравнений для равновесных режимов получается, если в (10.2) положить $dv/dt = 0$:

$$\begin{aligned}0 &= \Phi(v, f); \\ y &= \Psi(v, f).\end{aligned}\tag{10.3}$$

Получение СХ $y = F(f)$ сводится к исключению переменных v из уравнения (10.3), что в общем случае может быть непростой задачей.

Для частного случая системы второго порядка (см. систему уравнений (9.7)) развернутые уравнения (10.3) выглядят так:

$$\begin{aligned}\Phi_1(v_1, v_2, f) &= 0; \\ \Phi_2(v_1, v_2, f) &= 0; \\ y &= \Psi(v_1, v_2, f).\end{aligned}\tag{10.4}$$

Графическая иллюстрация определения $v_1(f^i)$ и $v_2(f^i)$ из первых двух уравнений системы (10.4) показана на рис. 10.1. Здесь f^i — i -е постоянное значение воздействия. Пересечения кривых, если решение существует, дают искомые значения

$$v_1^j(f^i), v_2^j(f^i),\tag{10.5}$$

где $j = 1, 2, 3$ — номер решения. Подставив (10.5) в третье уравнение системы (10.4), получим точки искомой СХ

$$y^j = \Psi(v_1^j(f^i), v_2^j(f^i)).$$

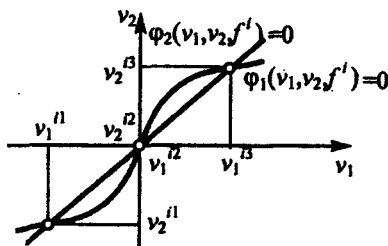


Рис. 10.1. Иллюстрация графической процедуры

Задавая другие значения f^i ; $i = 1, 2, \dots$, с выбранным шагом, получим множества точек, соединив (интерполяция) которых даст ветви $j = 1, 2, \dots$, в общем случае неоднозначной СХ системы.

В случае нескольких изолированных точек равновесия исследователь выделяет интересующие его точки и поочередно уточняет их координаты численным решением систем нелинейных уравнений.

10.3. ОПРЕДЕЛЕНИЕ РАВНОВЕСНЫХ РЕЖИМОВ И СТАТИЧЕСКИХ ХАРАКТЕРИСТИК СИСТЕМ С ТИПОВОЙ СТРУКТУРОЙ

Типовыми называют структуры систем, образованные последовательным и параллельным соединением звеньев, а также одноконтурные системы, образованные соединением двух звеньев с обратной связью.

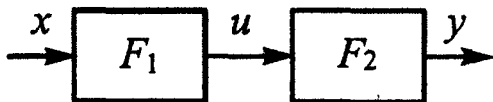


Рис. 10.2. Последовательное соединение НЭ

Эквивалентные СХ для последовательного соединения. Последовательно соединенные НЭ (рис. 10.2) эквивалентны одному НЭ со СХ

$$F_3 = F_2 \circ F_1,$$

являющейся композицией СХ составляющих:

$$y = F_2(u) = F_2(F_1(x)) = F_3(x).$$

В общем случае перестановка НЭ дает другой результат (нелинейные операторы не коммутативны), т. е.

$$F_2 \circ F_1 \neq F_1 \circ F_2.$$

Рассмотрим пример последовательного соединения идеального реле (см. рис. 9.6, а) и линейного безынерционного звена с коэффициентом усиления $k > 0$. Если линейный усилитель предшествует реле, то СХ соединения в точности повторяет СХ реле

$$y = C \operatorname{sgn} x;$$

если звенья переставить, то

$$y = kC \operatorname{sgn} x.$$

Бывают исключения, когда нелинейные преобразования коммутативны. Так, в случае *взаимно обратных* НЭ $F_1, F_2 = F_1^{-1}$ имеем:

$$y = F_1 \circ F_2(x) = F_2 \circ F_1(x) = x.$$

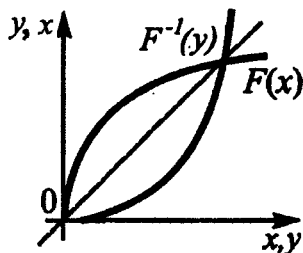


Рис. 10.3. Взаимнообратные СХ.

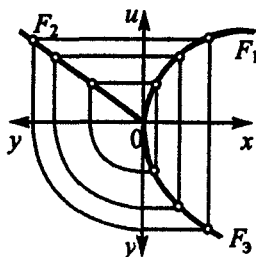


Рис. 10.4. Графическая процедура построения СХ последовательного соединения НЭ

Обратные характеристики F_1^{-1} ; F_2^{-1} существуют не всегда или не всегда однозначны при однозначной исходной. Взаимно обратные НЭ имеют СХ, графики которых являются симметричными относительно биссектрисы первого и третьего квадрантов (при одинаковом масштабе для абсцисс и ординат) (рис. 10.3).

В том случае, когда СХ последовательно соединенных НЭ заданы графически, легко получить СХ эквивалентного НЭ (рис. 10.4).

Эквивалентные СХ для параллельного соединения звеньев. Параллельно соединенные НЭ эквивалентны одному НЭ со СХ

$$y = F_1(x) + F_2(x) = F_3(x).$$

Здесь перестановка НЭ дает тот же результат.

Графическая процедура построения СХ $F_3(x)$ сводится к поординатному сложению СХ $F_1(x)$ и $F_2(x)$. Если

$$y = F_1(x) + F_2(x) = x,$$

то такие НЭ называют *взаимно дополнительными*. Компенсация естественной нелинейности достигается при параллельном подключении взаимно дополнительной нелинейности. Например, зона нечувствительности (см. рис. 9.7, б) может компенсироваться параллельным включением НЭ типа насыщение (см. рис. 9.7, а), зона линейности которого равна ширине зоны нечувствительности.

Если структурная схема не содержит контуров, т. е. образована только последовательным и параллельным соединениями звеньев, то легко найти постоянное значение выхода при заданном значении входа.

Эквивалентные СХ систем с обратной связью. Пусть два НЭ образуют контур — систему с отрицательной обратной связью (рис. 10.5, а). Принципиальная трудность определения значения выхода системы

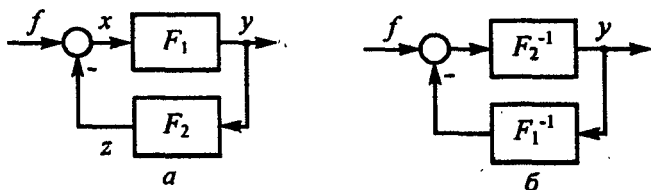


Рис. 10.5. Соединение звеньев:
а — с обратной связью; б — эквивалентное соединение

у при заданном значении входа f обусловлена обратной связью — для определения y следует знать значение x , зависящее от искомого y .

Для получения эквивалентной СХ необходимо исключить переменные x и z . Запишем выражение для выхода

$$y = F_1(x) = F_1(f - z) = F_1(f - F_2(y)). \quad (10.6)$$

Получена искомая зависимость, но в неявной форме.

Пусть существует обратная нелинейность F_1^{-1} . Применим это преобразование к левой и правой частям (10.6) и получим

$$f = F_1^{-1}(y) + F_2(y) = F_3^{-1}(y), \quad (10.7)$$

т. е. СХ, обратную искомой. Отсюда следует алгоритм графического построения искомой СХ (в предположении о ее обратимости):

- строится обратная СХ F_1^{-1} ;
- характеристики F_1^{-1} и F_2 суммируются;
- определяется СХ, обратная сумме.

Ниже будет показан пример построения эквивалентной СХ замкнутой системы $y = F_3(f)$ с использованием этого алгоритма.

Запишем выражение (10.7) в следующем порядке:

$$F_2(y) = f - F_1^{-1}(y). \quad (10.8)$$

Пусть существует обратная нелинейность F_2^{-1} . Применим соответствующее преобразование к обеим частям (10.8) и получим:

$$y = F_2^{-1}(f - F_1^{-1}(y)).$$

Этому выражению соответствует соединение с обратной связью двух НЭ (рис. 10.5, б), эквивалентное исходному соединению в смысле СХ $F_3(f)$. Сравнивая эту структуру с исходной (см. рис. 10.5, а), замечаем, что НЭ поменялись местами, а их характеристики заменены на обратные. На практике такой прием иногда используется для упрощения реализации НЭ [72].

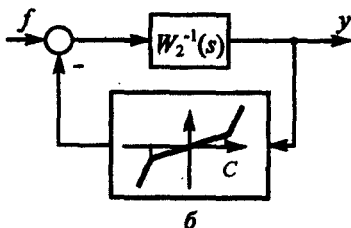
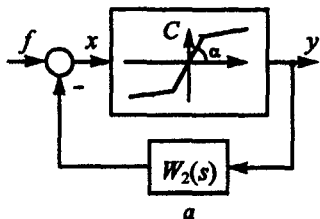


Рис. 10.6. Пример охвата НЭ линейным динамическим звеном

Эквивалентность структур, показанных на рис. 10.5, имеет место и для динамических звеньев. В случае линейных звеньев с передаточными функциями $W_1(s)$ и $W_2(s)$ обратными будут операторы $1/W_1(s)$ и $1/W_2(s)$ соответственно.

Рассмотрим пример (рис. 10.6, а), когда НЭ типа «насыщение» со СХ $F_1(x)$ охвачен отрицательной обратной связью в виде динамического звена с передаточной функцией (гибкая обратная связь)

$$W_2(s) = \tau s / (Ts + 1).$$

Эквивалентная структура показана на рис. 10.6, б. Нетрудно заметить, что в случае большого усиления НЭ в линейной зоне ($k = \operatorname{tg}(\alpha) \gg 1$), когда

$$|y(t)| < C,$$

можно в схеме рис. 10.6, б игнорировать обратную связь с коэффициентом передачи $1/k \ll 1$. В результате приближенно получим:

$$Y(s)/F(s) \approx (Ts + 1) / \tau s,$$

т. е. передаточную функцию пропорционально-интегрального звена. Такой способ часто применяется для реализации аналоговых регуляторов, реализующих ПИ-закон.

Ясно также, что СХ системы, образованной охваченным обратной связью интегратором, равна взаимнообратной СХ звена обратной связи, так как интегратор имеет бесконечное усиление по постоянному сигналу.

10.4. ОПРЕДЕЛЕНИЕ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК СИСТЕМ С ПРОИЗВОЛЬНОЙ СТРУКТУРОЙ

Пусть нелинейная модель системы управления представлена в форме структурной схемы. В ряде случаев СХ систем с произвольной структурой можно построить последовательным применением процедур, опи-

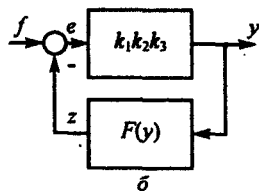
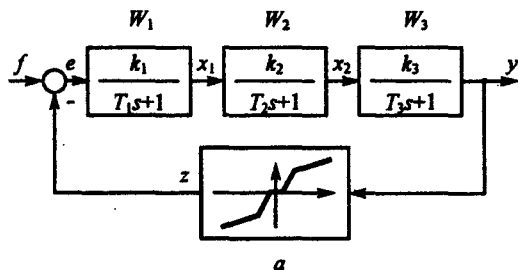


Рис. 10.7. Структурные схемы:
а — для динамических режимов; б — для статических режимов

санных в 10.3. Для этого вначале переходят к структурной схеме для равновесных режимов (постоянных сигналов), в которой статические линейные звенья описываются коэффициентами усиления, а НЭ — аналитическими выражениями или графиками СХ.

Покажем этот метод построения СХ на примере системы, структура которой представлена на рис. 10.7, а. Если все линейные динамические звенья заменить безынерционными звеньями с коэффициентами усиления

$$k_i = W_i(0); \quad i = 1, 2, 3,$$

то получим структурную схему для постоянных сигналов (рис. 10.7, б).

На рис. 10.8 показана процедура построения эквивалентной СХ замкнутой системы $y = F_3(f)$ с помощью алгоритма, полученного в § 10.3.

Выделим два особых случая: 1 — в системе имеются астатические звенья (интеграторы); 2 — в системе есть дифференцирующие звенья.

Для астатических звеньев с передаточными функциями:

$$W_i(s) = W'_i(s)/s; \quad W'_i(0) < \infty,$$

в равновесном режиме переменные входа равняются нулю. Это замечание упрощает решение задачи. Пусть в структурной схеме (см. рис. 10.7, а) передаточная функция третьего звена $W_3(s)$ имеет нулевой полюс. Тогда в равновесном режиме: $x_2 = 0$; $x_1 = 0$; $e = 0$; $z = f$, т. е. получаем:

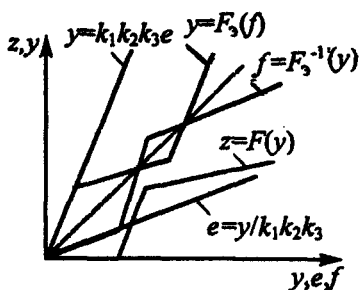


Рис. 10.8. Построение эквивалентной СХ системы с отрицательной обратной связью

$$z = F(y) = f.$$

Следовательно, СХ, обратная искомой, совпадает со СХ нелинейного элемента в обратной связи:

$$f = F(y) = F_s^{-1}(y).$$

Интегратор в прямой цепи за счет бесконечного усиления на нулевой частоте «переворачивает» СХ НЭ в цепи обратной связи.

Если в системе есть дифференцирующее звено

$$W_i(s) = sW_i'(s); \quad W_i'(0) \neq 0,$$

то в равновесном режиме его выход равен нулю. Например, наличие нулевого нуля хотя бы у одной из передаточных функций в схеме рис. 10.7, а приводит к тому, что связь по постоянному сигналу прерывается, т. е. для равновесного режима $y = 0$.

Следует отметить, что, как и в случае линейных систем (см. п. 2.11.1), не для любых графов существует последовательность эквивалентных преобразований, приводящая к цели. В линейных системах задача полного сворачивания таких графов решалась дополнительным применением правил переноса точек съема и суммирования сигналов. В нелинейных системах правила переноса в общем случае не выполняются. Для построения СХ нелинейных систем здесь следует записывать системы уравнений и обращаться к численным процедурам их решения.

10.5. РАВНОВЕСНЫЕ РЕЖИМЫ В СИСТЕМАХ УПРАВЛЕНИЯ И КАТАСТРОФЫ

Пусть система находится в равновесном состоянии. Если ее параметры медленно изменяются, то естественно предположить, что положение равновесия также будет изменяться. В том случае, когда при некотором значении параметров рассматриваемое положение равновесия перестанет существовать, система будет вынуждена перескочить в другое устойчивое равновесное состояние, которое может быть весьма далеким от первоначального. Таким образом, эволюция системы завершается скачкообразным изменением или, как говорят, *катастрофой* (например, разрушение экономики или строительного сооружения).

В линейных системах положение равновесия единственно. Рассмотрим простейший пример системы с обратной связью, на вход которой

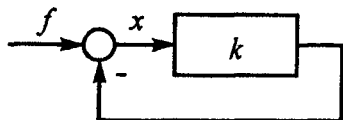


Рис. 10.9. Линейная система

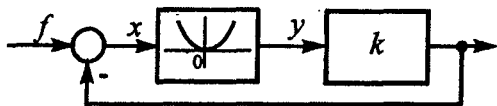


Рис. 10.10. Пример нелинейной системы

подано воздействие f (рис. 10.9). При постоянном значении f значение переменной x равно $x = f / (1 + k)$.

Если усиление изменяется ($k \rightarrow -1$), то положение равновесия становится большим положительным числом и уходит в $+\infty$. При $k = -1 - \varepsilon$ ($\varepsilon > 0$, малое) координата положения равновесия принимает большие по модулю отрицательные значения. Говорят, что значение $k = -1$ соответствует катастрофе.

Нелинейные системы могут иметь несколько положений равновесия. Можно ожидать, что при рассмотрении эволюций нелинейных систем окажется много вариантов возникновения устойчивых и неустойчивых положений равновесия. Однако показано [9], что для широко распространенного вида динамических систем, в которых на практике наблюдаются скачкообразные изменения, можно дать геометрическое описание путей, по которым происходят такие изменения. Эти пути принадлежат небольшому числу классифицируемых типов, названных *элементарными катастрофами*.

Ниже рассмотрим примеры нелинейных систем управления, в которых наблюдаются некоторые типы элементарных катастроф.

Пример 1. Катастрофы типа «складка». Пусть НЭ системы расчетной структуры (рис. 10.10) является квадрататором

$$y = F(x) \equiv x^2.$$

На вход системы подано постоянное воздействие f , а коэффициент усиления ЛЧ равен k . Переменные системы определяются двумя параметрами — уровнем воздействия f и коэффициентом усиления k , которые могут медленно изменяться. Найдём значения переменной x на входе НЭ в равновесных состояниях и изучим, как влияют на эти значения изменения параметров k и f .

Для решения задачи запишем уравнения статики:

$$\begin{aligned} y &= x^2; \\ y &= (f - x)/k. \end{aligned} \quad (10.9)$$

Графическое определение положений равновесия иллюстрирует рис. 10.11, а. В зависимости от соотношения параметров k и f система имеет два положения равновесия (прямая 1), одно (прямая 2) или ни одного (прямая 3). Соотношение параметров (k, f) , отвечающее касанию прямой и параболы, является критическим — сколь

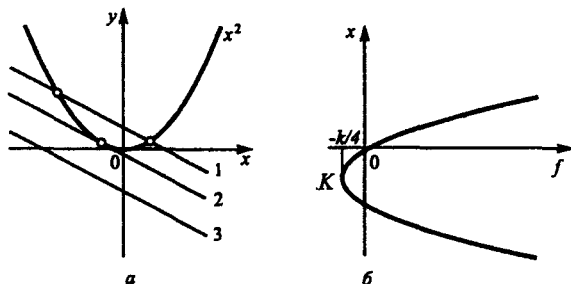


Рис. 10.11. Пример катастрофы типа «складка»:
 а – графическое определение положений равновесия;
 б – зависимость положения равновесия от параметра

угодно малое изменение параметров приводит к качественному изменению ситуации. Множество значений (k, f) , когда два положения равновесия сливаются в одно, называют *множеством катастроф K*.

Для аналитического определения множества K вместо системы (10.9) запишем одно уравнение для положений равновесия

$$kx^2 + x - f = 0.$$

Формула определения координаты x для положений равновесия

$$x_{1,2} = (-1 \pm \sqrt{1 + 4kf}) / 2$$

даст действительные значения при условии

$$4kf + 1 \geq 0, \quad f = -1/4k,$$

причем равенству отвечает критическая ситуация, когда два положения равновесия сливаются в одно. Множество катастроф K на плоскости (k, f) описывается уравнением

$$4kf + 1 = 0.$$

Если изменяется только один из параметров, например f , а $k > 0$, то график зависимости положений равновесия $x(f)$ имеет вид параболы (рис. 10.11, б). Ветви параболы при $f > -k/4$ соответствуют двум положениям равновесия. При $f = -k/4$ они сливаются в одно (точка катастрофы), а при дальнейшем уменьшении значения f положение равновесия исчезает (уходит в $-\infty$).

Данный тип катастроф получил название «складка».

Пример 2. Катастрофы типа «сборка». Пусть НЭ системы (см. рис. 10.10) имеет СХ, описываемую полиномом третьей степени

$$y = F(x) \equiv x^3 - x.$$

Для графической иллюстрации анализа равновесных режимов уравнения статики запишем так:

$$\begin{aligned} y &= x^3 - x, \\ y &= (f - x)/k. \end{aligned} \tag{10.10}$$

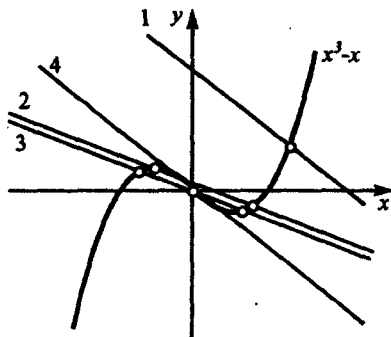


Рис. 10.12. Графическое определение положений равновесия

Определение положений равновесия иллюстрирует рис. 10.12. В зависимости от соотношения параметров k и f система имеет три положения равновесия (прямая 3), два (прямая 2) или одно (прямые 1 и 4). При этом соотношение параметров (f, k) , отвечающее касанию прямой и кривой СХ НЭ, является катастрофическим — сколь угодно малое количественное изменение параметров приводит к качественному изменению ситуации.

Для определения множества катастроф K — множества значений параметров, когда положения равновесия сливаются, запишем вместо системы (10.10) одно уравнение для положений равновесия

$$\varphi(x) \equiv kx^3 - (k-1)x - f = 0. \quad (10.11)$$

Действительным нулям функции $\varphi(x)$, т. е. пересечениям кривой $\varphi(x)$ оси абсцисс соответствуют положения равновесия (рис. 10.13). Рис. 10.13, а соответствует одному положению равновесия (прямая 1 на рис. 10.12). Катастрофы отвечают тем значениям параметров k и f , когда вместо пересечений имеет место касание кривой $\varphi(x)$ оси абсцисс (рис. 10.13, б и прямая 2 на рис. 10.12). При дальнейшем изменении параметров вместо трех положений равновесия (рис. 10.13, в и прямая 3 на рис. 10.12) имеем только одно (рис. 10.13, г и прямая 4 на рис. 10.12).

Таким образом, катастрофам отвечают те значения параметров, когда нули функции $\varphi(x)$ совпадают с ее стационарными точками, т. е. дополнительно выполняется условие:

$$\varphi'(x) \equiv 3kx^2 - k + 1. \quad (10.12)$$

Совместное решение уравнений (10.11) и (10.12) дает выражение для кривой катастроф K на плоскости (k, f) :

$$f = \pm \frac{2(k-1)}{3} \sqrt{\frac{k-1}{3k}}.$$

На рис. 10.14 изображена кривая катастроф K типа «сборка». В правой полуплоскости, т. е. для $k > 0$, выделяется полукубическая парабола с точкой возврата (острием) в точке $(k=1, f=0)$. Эта кривая делит полуплоскость на две части: меньшую и большую. Точки меньшей части соответствуют трем положениям равновесия, точки большей ча-

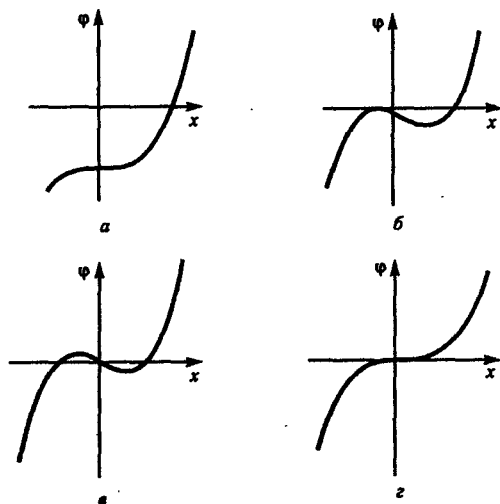


Рис.10.13. Поведение кривой $\varphi(x)$

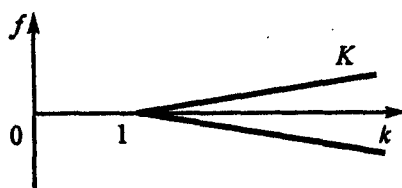


Рис.10.14. Кривые катастрофы типа «сборка»

сти — одному, а точки кривой — двум. При подходе к кривой K из меньшей части два положения равновесия (из трех) сливаются и исчезают. При подходе к острю сливаются все три положения равновесия (прямая 4 на рис. 10.12).

Глава 11

АНАЛИЗ ПОВЕДЕНИЯ СИСТЕМ НА ФАЗОВОЙ ПЛОСКОСТИ

11.1. МЕТОД ФАЗОВОЙ ПЛОСКОСТИ

Рассмотрим математическую модель системы в форме дифференциальных уравнений первого порядка, разрешенных относительно производных

$$\frac{dv_i}{dt} = \varphi_i(v_1, \dots, v_n); \quad i = 1, \dots, n. \quad (11.1)$$

В правые части уравнений (11.1) время t явно не входит, такие системы называют *автономными*. На автономные системы не действуют внешние силы. Автономные системы *стационарны* — их свойства неизменны во времени.

Состояние конечномерной динамической системы (11.1) характеризуется вектором в n -мерном пространстве $\mathbf{v}(t) = (v_1(t), \dots, v_n(t))^T$. Начальное состояние $\mathbf{v}(0)$ автономной системы (11.1) полностью определяет ее поведение для $t > 0$ независимо от предыстории, т. е. от того, каким путем система пришла в это состояние. Геометрическое место точек концов вектора $\mathbf{v}(t)$ при $t \geq 0$ образует *траекторию состояния* — образ поведения при конкретном начальном состоянии. На рис. 11.1 иллюстрируется траектория трехмерной системы.

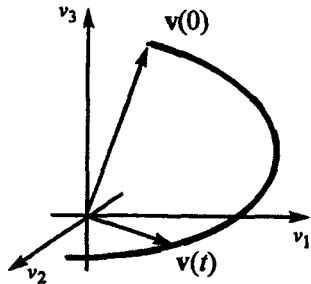


Рис. 11.1. Траектория изменения состояния

Несмотря на то что интерпретация метода пространства состояний распространяется на системы любого порядка, важное его преимущество — наглядность — наиболее ярко проявляется в случае систем второго порядка, когда состояния системы представляются точками на *фазовой плоскости*. Следует добавить, что нелинейные модели второго порядка позволяют выявлять многие принципиальные особенности поведения динамических систем; это определяет методическое, теоретическое и практическое значение метода фазовой плоскости. Применение метода может быть оправданным для предварительного анализа новой системы по упрощенным моделям, поскольку метод фазовой плоскости дает наглядную картину общего характера поведения рассматриваемой динамической системы. Во второй части анализа качественные исследования, основанные на таком представлении; следует дополнить количественными, с помощью которых можно получить численные результаты, но трудно выявить общие закономерности движения.

Метод фазовой плоскости дает возможность изобразить качественную картину всей совокупности свободных движений (процессов) для выбранной области начальных условий (состояний), а при необходимости — провести точные исследования интересующих типов движений.

Пусть заданы уравнения системы второго порядка

$$\begin{aligned} \frac{dv_1}{dt} &= \varphi_1(v_1, v_2); \\ \frac{dv_2}{dt} &= \varphi_2(v_1, v_2). \end{aligned} \quad (11.2)$$

Для получения фазовой траектории при начальных условиях $v_{10} = v_1(0)$ и $v_{20} = v_2(0)$ из решений системы (11.2)

$$v_1(t; v_{10}, v_{20}); \quad v_2(t; v_{10}, v_{20})$$

следует исключить время t :

$$v_2(v_1; v_{10}, v_{20}).$$

На рис. 11.2 показана графическая процедура исключения t и построения фазовой траектории.

Если изменить начальные условия, но так, что вектор начального состояния попадает на ранее построенную фазовую траекторию (точка α на рис. 11.2, б), то новая траектория совпадает со старой. Следовательно, отрезок фазовой траектории изображает бесчисленное множество движений при начальных состояниях системы, совпадающих с траекторией. Это упрощает вывод обобщающих суждений о свойствах системы по совокупностям фазовых траекторий — *фазовому портрету* динамической системы.

В случае автономности динамических систем время t явно не входит в правую часть дифференциальных уравнений (11.2). Рассматривая координату v_1 как независимую переменную, можно путем деления второго уравнения системы (11.2) на первое исключить время и получить одно дифференциальное уравнение первого порядка

$$\frac{dv_2}{dv_1} = \frac{\varphi_2(v_1, v_2)}{\varphi_1(v_1, v_2)}. \quad (11.3)$$

Решая это уравнение при конкретных начальных условиях, можно получить фазовые траектории.

Если функции $\varphi_1(v_1, v_2)$ и $\varphi_2(v_1, v_2)$ однозначны, то каждой точке (v_1, v_2) соответствует единственное значение производной dv_2/dv_1 (наклона касательной к фазовой траектории), т. е. через эту точку фазовой плоско-

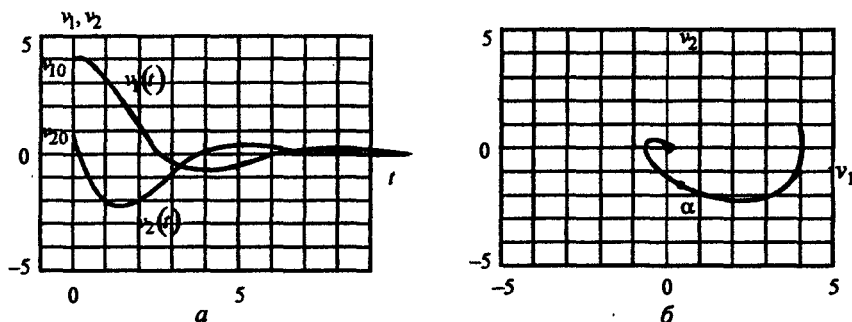


Рис. 11.2. Графическая процедура построения фазовой траектории

сти проходит только одна фазовая траектория. Единственность нарушается в так называемых *особых точках*, соответствующих состояниям равновесия системы:

$$\begin{aligned} dv_1/dt &= 0; \\ dv_2/dt &= 0, \end{aligned} \quad (11.4)$$

координаты которых определяются из уравнений

$$\begin{aligned} \varphi_1(v_1, v_2) &= 0; \\ \varphi_2(v_1, v_2) &= 0. \end{aligned} \quad (11.5)$$

В особых точках, как это следует из (11.3) и (11.4), имеется неопределенность производной: $dv_2/dv_1 = 0/0$. Каждая особая точка изображает отдельное (тривиальное) решение системы (11.2) и должна рассматриваться как отдельная фазовая траектория.

В качестве фазовой переменной v_1 часто выбирается переменная входа нелинейного элемента; $v_1 = x$ (см. рис. 9.16). Переменная v_2 обычно выбирается из условия $v_2 = dv_1/dt = x'$. В этом случае так называемого естественного базиса система уравнений (11.2) примет вид:

$$\begin{aligned} dx/dt &= x'; \\ dx'/dt &= \varphi_2(x, x'), \end{aligned}$$

а вместо (11.3) имеем:

$$\frac{dx'}{dx} = \frac{\varphi_2(x, x')}{x'}. \quad (11.6)$$

Выполнение условия $v_2 = dv_1/dt$ придает фазовой плоскости следующие свойства:

- особые точки располагаются только на оси абсцисс, где $v_2 = dv_1/dt = 0$;
- в верхней полуплоскости ($v_2 = dv_1/dt > 0$) фазовые траектории направлены слева направо, т. е. в сторону увеличения v_1 , а в нижней полуплоскости — справа налево;
- фазовые траектории ортогональны оси абсцисс, так как из (11.6) при $v_2 = 0$ следует $dv_2/dv_1 = \infty$.

В заключение заметим, что для систем первого порядка ($n = 1$) дифференциальное уравнение системы

$$x' = \varphi(x)$$

одновременно описывает фазовую траекторию в естественном базисе (x, x') . Положениями равновесия являются действительные нули функции $\varphi(x)$, а фазовый портрет вырождается в единственную траекторию.

11.2. ПОВЕДЕНИЕ НЕЛИНЕЙНЫХ СИСТЕМ В ОКРЕСТНОСТИ ПОЛОЖЕНИЙ РАВНОВЕСИЯ

Рассмотрим i -е изолированное положение равновесия (v_1^i, v_2^i) , т. е. i -ю особую точку, являющуюся решением системы уравнений (11.5). Пусть функции φ_1 и φ_2 в окрестности этой точки допускают разложение в степенной ряд по отклонениям переменных $(\Delta v_1, \Delta v_2)$. Пренебрегая нелинейными членами разложения, вместо нелинейных уравнений (11.2) получим систему линейных дифференциальных уравнений для малых отклонений переменных:

$$\begin{aligned}\frac{d\Delta v_1}{dt} &= a_{11}\Delta v_1 + a_{12}\Delta v_2; \\ \frac{d\Delta v_2}{dt} &= a_{21}\Delta v_1 + a_{22}\Delta v_2.\end{aligned}\quad (11.7)$$

В матричной форме уравнения (11.7) запишутся так:

$$\frac{d\Delta v}{dt} = A\Delta v, \quad (11.8)$$

где

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \frac{d\varphi_1}{dv_1} & \frac{d\varphi_1}{dv_2} \\ \frac{d\varphi_2}{dv_1} & \frac{d\varphi_2}{dv_2} \end{pmatrix}$$

— матрица Якоби системы (11.2) в рассматриваемой точке равновесия.

Процедура переноса начала координат в особую точку и линейризации правых частей (11.2) показана на рис. 11.3. Касательные к кривым пересекаются, если якобиан отличен от нуля:

$$\det A \neq 0.$$

Решения уравнений (11.7) имеют вид:

$$\begin{aligned}\Delta v_1(t) &= C_{11}e^{s_1 t} + C_{12}e^{s_2 t}; \\ \Delta v_2(t) &= \dot{C}_{21}e^{s_1 t} + C_{22}e^{s_2 t},\end{aligned}\quad (11.9)$$

где постоянные C_{ij} зависят от начальных условий $\Delta v_1(0)$, $\Delta v_2(0)$ и корней характеристического полинома

$$D(s) = \det(sI - A). \quad (11.10)$$

Выражения (11.9) соответствуют случаю некротных корней s_1, s_2 .

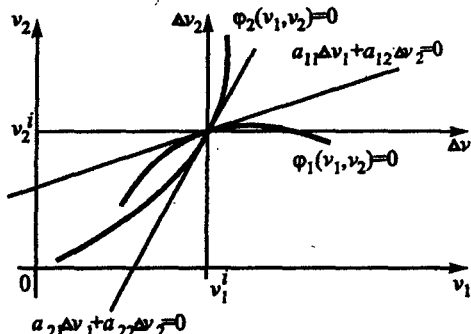


Рис. 11.3. Линеаризация в окрестности положения равновесия

Оказывается, что фазовые портреты нелинейных систем в окрестности особых точек близки к фазовым портретам линеаризованных систем (11.8). Рассмотрим типы особых точек линейных систем второго порядка в зависимости от расположения корней характеристического полинома (11.10), т. е. собственных значений матрицы A .

Если корни s_1, s_2 полинома (11.10) действительные отрицательные, то имеет место особая точка типа «устойчивый узел» (рис. 11.4, а). На рис. 11.4, б показаны кривые $v_1(t)$ для двух начальных состояний. Прямолinéйнým фазовым траекториям соответствуют состояния, когда постоянные при одной из экспонент (11.9) равняются нулю ($C_{11} = C_{21} = 0$ или $C_{12} = C_{22} = 0$).

При положительных действительных корнях s_1, s_2 имеем особую точку типа «неустойчивый узел» (рис. 11.5, а). На рис. 11.5, б показаны примеры соответствующих кривых изменения переменной v_1 во времени.

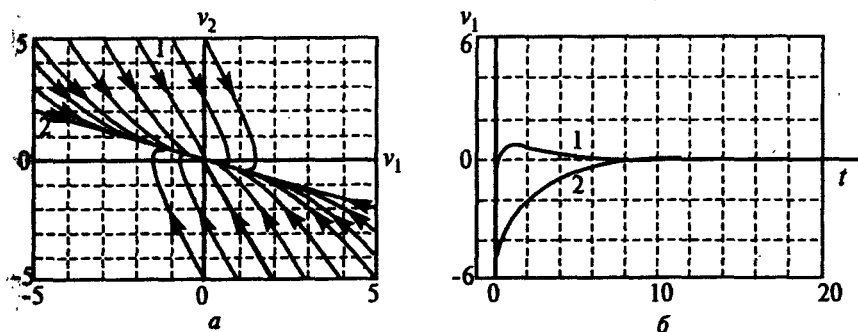


Рис. 11.4. Особая точка типа «устойчивый узел»

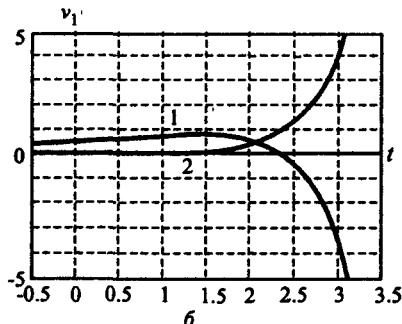
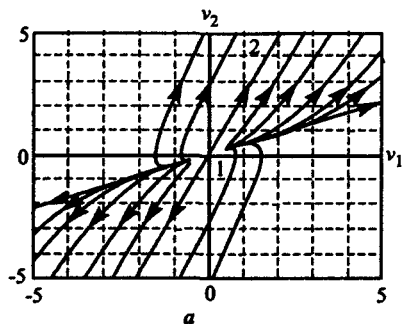


Рис. 11.5. Особая точка типа «неустойчивый узел»

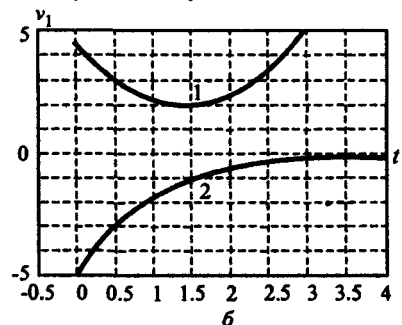
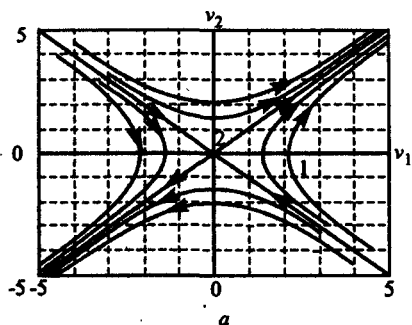


Рис. 11.6. Особая точка типа «седло»

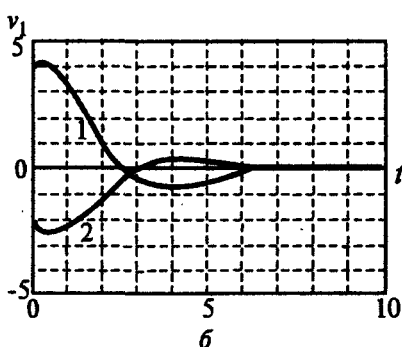
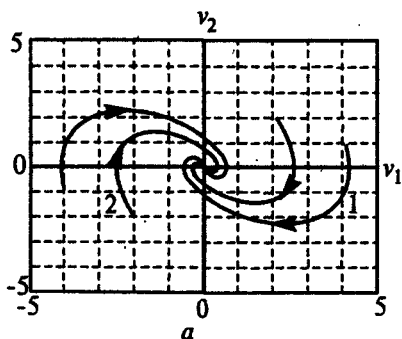


Рис. 11.7. Особая точка типа «устойчивый фокус»

Если один из действительных корней отрицательный, а другой — положительный, то особая точка называется «седло» (рис.11.6, а). Прямолинейные траектории — сепаратрисы седла — соответствуют случаю, когда один из коэффициентов в (11.9) равен нулю (прямая 2).

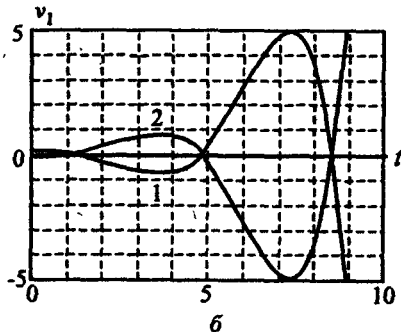
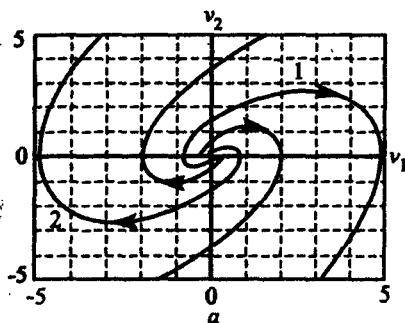


Рис. 11.8. Особая точка типа «неустойчивый фокус»

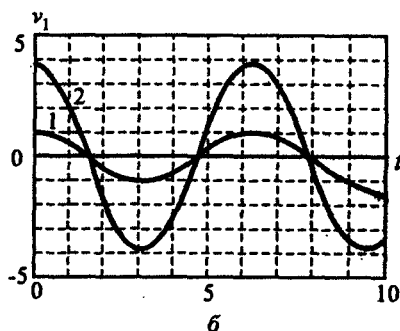
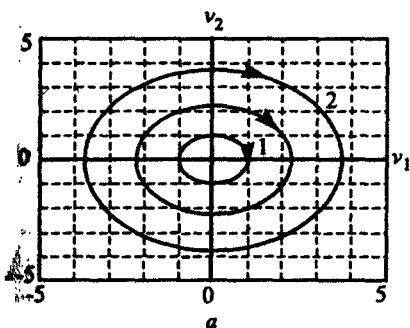


Рис. 11.9. Особая точка типа «центр»

Комплексно-сопряженные корни с отрицательными действительными частями дают особую точку типа «устойчивый фокус» (рис. 11.7, а). На рис. 11.7, б показан соответствующий затухающий колебательный процесс.

Если действительные части комплексных корней положительны, то имеем «неустойчивый фокус» (рис. 11.8, а) и колебательные расходящиеся процессы (рис. 11.8, б).

Мнимым корням консервативной системы $s_{1,2} = \pm j\omega_0$ соответствует особая точка типа «центр», образованная вложенными друг в друга эллипсами (рис. 11.9, а). Незатухающий колебательный процесс гармонической формы с периодом $T = 2\pi/\omega_0$ имеет амплитуду, определяемую начальными условиями (рис. 11.9, б).

Особым точкам типа «устойчивый узел» и «устойчивый фокус» соответствуют устойчивые «в малом» положения равновесия нелинейной системы; в случае особых точек «седло», «неустойчивый узел» и «не-

устойчивый фокус» эти положения равновесия неустойчивы. В случае особой точки типа «центр» об устойчивости положения равновесия нелинейной системы нельзя судить по линеаризованным уравнениям (11.8).

11.3. ПОСТРОЕНИЕ ФАЗОВЫХ ПОРТРЕТОВ НЕЛИНЕЙНЫХ СИСТЕМ

Исследование поведения нелинейных систем второго порядка на фазовой плоскости целесообразно начинать с определения положений равновесия — особых точек фазового портрета. Принципиальным отличием нелинейных систем является то, что положений равновесия может быть несколько или они могут отсутствовать вообще.

Если в окрестности изолированных положений равновесия нелинейности достаточно гладкие, то записывают линеаризованные дифференциальные уравнения для малых отклонений от конкретных положений равновесия. По расположению собственных значений матриц или корней характеристического полинома устанавливаются типы особых точек (см. § 11.2). В окрестности изолированных особых точек — положений равновесия — фазовые портреты нелинейных систем похожи на фазовые портреты соответствующих линеаризованных систем. По мере удаления от точек равновесия фазовые портреты нелинейных систем могут иметь качественные отличия.

Для построения фазовых траекторий используют графические, аналитические и численные методы.

11.3.1. Метод изоклин

Наиболее простым графическим методом решения дифференциальных уравнений (11.3), (11.6) является метод изоклин. *Изоклина* — кривая равного наклона касательных к фазовым траекториям. Ее уравнение получается приравниванием правых частей (11.3) или (11.6) заданному наклону c :

$$\Phi_2(v_1, v_2) / \Phi_1(v_1, v_2) = c$$

или

$$\Phi_2(v_1, v_2) / v_2 = c.$$

Параметр c семейства изоклин равен тангенсу угла наклона касательных к фазовым траекториям (при одинаковых масштабах обеих осей координат фазовой плоскости). Совокупность наклонов c , выбирается

так, чтобы изменение угла между двумя соседними изоклинами было примерно одинаковым.

11.3.2. Примеры аналитического построения фазовых портретов

В некоторых случаях возможно аналитическое решение дифференциальных уравнений для фазовых траекторий (11.3), (11.6).

Пример 1. Рассмотрим уравнение математического маятника

$$y'' - \omega_0^2 \sin y = 0.$$

Как указывалось в § 9.1, маятник имеет счетное множество положений равновесия: $y = k\pi$; $k = 0, \pm 1, \pm 2, \dots$. Здесь за начальное положение маятника принято верхнее.

Линеаризованное уравнение для малых отклонений от нижних положений равновесия (нечетных $k = \pm 1, \pm 3, \dots$) запишем в виде

$$\Delta y'' + \omega_0^2 \Delta y = 0,$$

для отклонений от верхних положений равновесия (четных $k = 0, \pm 2, \dots$) — в виде:

$$\Delta y'' - \omega_0^2 \Delta y = 0.$$

Характеристический полином дифференциального уравнения для нижних положений равновесия

$$D(s) = s^2 + \omega_0^2$$

имеет пару мнимых корней $s_{1,2} = \pm j\omega_0$. Следовательно, этим положениям равновесия отвечает особая точка типа «центр» (см. рис. 11.9). Характеристический полином для верхних положений равновесия

$$D(s) = s^2 - \omega_0^2$$

имеет действительные корни разных знаков $s_{1,2} = \pm \omega_0$; этим положениям равновесия отвечает особая точка типа «седло» (см. рис. 11.6).

Для получения аналитических выражений для фазовых траекторий маятника запишем его дифференциальное уравнение в форме Коши, приняв следующие переменные состояния: $v_1 = y$; $v_2 = y'$ (естественный базис):

$$\begin{aligned} \frac{dy}{dt} &= y'; \\ \frac{dy'}{dt} &= \omega_0^2 \sin y. \end{aligned}$$

Деление второго уравнения на первое даст дифференциальное уравнение первого порядка:

$$\frac{dy'}{dy} = \omega_0^2 \frac{\sin y}{y'}, \quad (11.11)$$

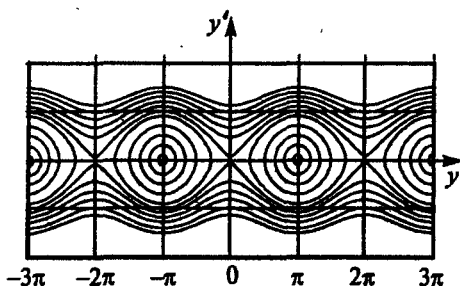


Рис. 11.10. Фазовый портрет маятника

решая которое при различных начальных условиях (y_0, y'_0) получаем выражения для интегральных кривых на плоскости (y, y') . Уравнение (11.11) допускает разделение переменных

$$y' dy' = \omega_0^2 \sin y dy;$$

после интегрирования имеем

$$\frac{(y')^2}{2} \Big|_{y'_0}^{y'} = -\omega_0^2 \cos y \Big|_{y_0}^y$$

или искомое выражение для фазовых траекторий

$$y' = \pm \sqrt{2\omega_0^2 (\cos y_0 - \cos y) + (y'_0)^2}.$$

Соответствующий фазовый портрет маятника изображен на рис. 11.10. *Сепаратрисы*, образованные слиянием искривленных «усов» седел, оказываются границами областей с колебательным движением; незамкнутые траектории вне этих областей изображают неравномерные вращательные движения маятника вокруг точки подвеса.

Пример 2. Фазовый портрет релейной системы (рис. 11.11).

В случае НЭ типа «реле с зоной нечувствительности» можно сразу найти положения равновесия: $y = 0$ как вход интегратора в равновесном состоянии; $|x| \leq b$; $|z| \leq b$. Имеет место отрезок равновесия.

Для построения фазового портрета запишем дифференциальные уравнения системы в форме Коши; выберем следующие переменные состояния: x, x' . Дифференциальное уравнение линейной части

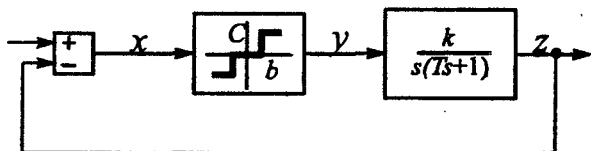


Рис. 11.11. Релейная система

$$Td^2z/dt^2 + dz/dt = ky$$

с учетом $z = -x$ запишется так:

$$Tdx'/dt + x' = -ky.$$

Отсюда имеем искомую форму уравнений:

$$dx/dt = x';$$

$$dx'/dt = -(x' + ky)/T = -(x' + kF(x))/T.$$

Поделим второе уравнение на первое:

$$dx'/dx = -(x' + kF(x))/Tx'. \quad (11.12)$$

Если учесть, что на различных интервалах x функция $F(x)$ постоянна, то в уравнении (11.12) разделяются переменные, т. е. имеем:

$$\frac{x'dx'}{x' + kF(x)} = -\frac{1}{T}dx.$$

Интегрирование последнего уравнения дает

$$x' - kF(x) \ln|x' + kF(x)| = -x/T + R,$$

где R — постоянная, определяемая начальными условиями. Для различных интервалов x получим следующие уравнения для фазовых траекторий:

$$x < -b: \quad x' + kC \ln|x' - kC| = -x/T + R; \quad (11.13)$$

$$|x| \leq b: \quad x' = -x/T + R;$$

$$x > b: \quad x' - kC \ln|x' + kC| = -x/T + R. \quad (11.14)$$

На рис. 11.12 показан фазовый портрет релейной системы «с зоной нечувствительности». Линиям $x = \pm b$ (так называемым линиям переключения реле) соответствуют границы трех «листов» фазовой плоскости. Движения системы завершаются на отрезке равновесия.

Если модель системы содержит НЭ с кусочно-постоянной (или кусочно-линейной) характеристикой, то процесс может быть разбит на ряд интервалов так, что в

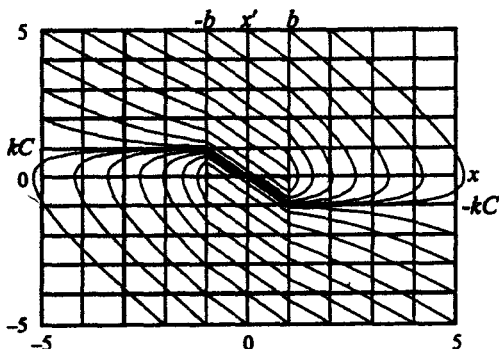


Рис. 11.12. Фазовый портрет в случае НЭ типа «реле с зоной нечувствительности»

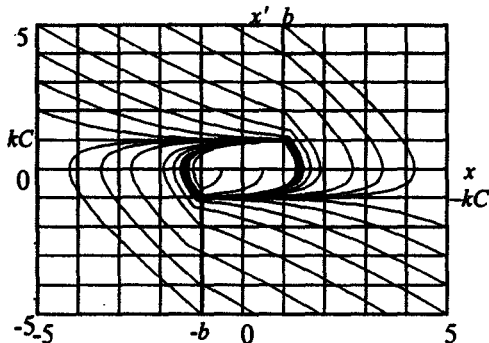


Рис. 11.13. Фазовый портрет в случае НЭ типа «реле с гистерезисом»

пределах каждого интервала движение описывается линейными дифференциальными уравнениями. На фазовой плоскости каждому линейному участку такой характеристики нелинейного элемента соответствует отдельная область или лист, в пределах которого правые части дифференциальных уравнений (11.2) линейны, а фазовые траектории составлены из дуг траекторий линейных систем. Излому или разрыву таких СХ НЭ соответствует граница листа фазовой плоскости (линия переключения).

Пример 3. Изменим условия Примера 2: пусть НЭ имеет характеристику типа «реле с гистерезисом» (см. рис. 9.6, б). В этом случае система не может иметь положений равновесия, так как всегда $y \neq 0$. Фазовая плоскость состоит из двух листов, на каждом из них траектории описываются выражениями (11.13) или (11.14). Границами листов являются в верхней полуплоскости прямая $x = b$, а в нижней — $x = -b$. Фазовый портрет показан на рис. 11.13. Выделяется устойчивый предельный цикл, значит система функционирует в автоколебательном режиме.

11.4. СВЯЗЬ ФАЗОВЫХ ТРАЕКТОРИЙ СО ВРЕМЕНЕМ

Как показано в § 11.1, построение фазовых траекторий связано с исключением времени t из решений $v_1(t; v_{10}, v_{20})$ и $v_2(t; v_{10}, v_{20})$. Представляет интерес обратный переход — построение процессов по интересующему отрезку фазовой траектории. По существу, речь идет об определении времени «пробегания» изображающей точкой отрезков фазовой траектории — проставлении на траекториях меток времени.

Ограничимся рассмотрением случая естественного базиса на фазовой плоскости: $v_1 = x$; $v_2 = x'$. Зафиксируем интересующий нас отрезок фазовой траектории $x'(x)$, как это показано на рис. 11.14. Найдем, за какое время Δt изображающая точка пройдет от начала отрезка ($x = x_0$) до его конца ($x = x_f$).

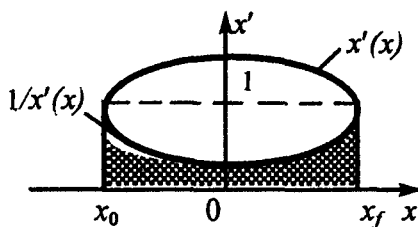


Рис. 11.14. Определение длительности процесса

По определению имеем

$$dx/dt = x';$$

отсюда следует

$$dx/x' = dt.$$

Интегрируя, получим

$$\int_{x_0}^{x_f} \frac{dx}{x'(x)} = t \Big|_{t_0}^{t_f} = t_f - t_0 = \Delta t.$$

Результат имеет простой геометрический смысл — время процесса (перехода системы из одного состояния в другое) пропорционально площади под кривой $1/x'(x)$. На рис. 11.14 показана эта кривая; площадь под ней заштрихована.

Чем выше проходит траектория $x'(x)$ в верхней полуплоскости, а значит, чем больше скорости процесса, тем ближе кривая $1/x'(x)$ к оси абсцисс, тем меньше Δt , т. е. быстрее протекает процесс. На рис. 11.15, а показаны отрезки траекторий двух различных систем между одними и теми же точками. Ясно, что процесс в первой системе (кривая 1) протекает быстрее по сравнению со второй системой (кривая 2) (рис. 11.15, б).

Рассмотрим пример фазовой траектории в виде отрезка прямой, направленной к началу координат (рис. 11.16, а)

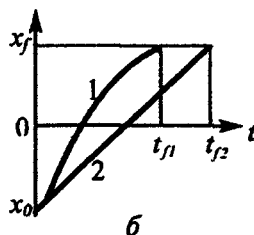
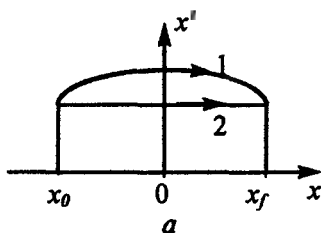


Рис. 11.15. Фазовые траектории и процессы

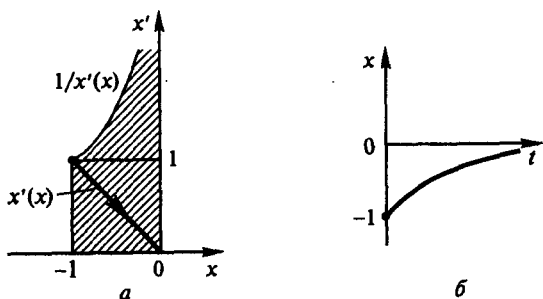


Рис. 11.16. Фазовая траектория и процесс

$$dx/dt = x' = -x. \quad (11.15)$$

Обратная кривая $1/x'(x) = -1/x$ — кусок гиперболы; площадь под ней равна времени процесса — прихода изображающей точки в начало координат. Найдем это время:

$$\Delta t = - \int_{-1}^0 dx/x = - \ln x \Big|_{-1}^0 = \infty.$$

Действительно, прямолинейному отрезку фазовой траектории соответствует экспоненциальное движение (рис. 11.16, б), т. е. решение дифференциального уравнения (11.15) при начальном условии $x(0) = -1$ имеет вид

$$x(t) = -e^{-t}.$$

Такое движение затухает бесконечно долго.

11.5. ОСОБЕННОСТИ ФАЗОВЫХ ПОРТРЕТОВ НЕЛИНЕЙНЫХ СИСТЕМ

В линейной системе всегда имеется только одно положение равновесия и соответственно только одна особая точка пространства состояний (фазового пространства). Область притяжения этой особой точки, если система устойчива, совпадает со всем фазовым пространством. Фазовое пространство линейных систем не содержит изолированные замкнутые траектории (предельные циклы).

В окрестности изолированных положений равновесия фазовые портреты нелинейных систем с гладкими в этих точках характеристиками подобны портретам соответствующих линейных систем. Но по мере удаления от особой точки фазовый портрет нелинейной системы может иметь существенные отличия. Рассмотрим основные из них.

Неединственность особых точек. Качественное отличие фазовых портретов нелинейных систем от линейных может быть обусловлено наличием множества особых точек различных типов. В примере фазового портрета маятника (см. рис. 11.10) можно видеть счетное множество особых точек двух типов — «центр» и «седло». Возможно существование бесконечного числа точек равновесия в виде отрезка равновесия; это имело место в примере фазового портрета системы, НЭ которой представляет собой реле с зоной нечувствительности (см. рис. 11.12). Наконец, фазовые портреты некоторых нелинейных систем вообще не имеют особых точек. Примером является портрет нелинейной системы с гистерезисной релейной характеристикой (см. рис. 11.13).

Предельные циклы и сепаратрисы. На фазовых портретах некоторых нелинейных систем имеют место изолированные замкнутые траектории — *предельные циклы* Пуанкаре (H. Poincaré). Если предельный цикл устойчив, то ему отвечают автоколебания в системе. Причем, эти колебания устойчивы не только по отношению к малым изменениям начальных условий, но также и к малым изменениям параметров системы. На рис. 11.17, а приведена иллюстрация фазового портрета некоторой нелинейной системы с устойчивым предельным циклом.

В качестве другого примера приведем фазовый портрет осциллятора Ван дер Поля с устойчивым предельным циклом (рис. 11.18).

В том случае, когда предельный цикл неустойчив по начальным условиям, в реальной системе нельзя наблюдать установившиеся колебания, так как они расходятся или затухают из-за малых возмущений. Устойчивые и неустойчивые предельные циклы являются границами областей фазовой плоскости с различным характером движения. Например, неустойчивый внутренний предельный цикл фазового портрета, изображенного на рис. 11.17, б, ограничивает открытую область притяжения устойчивой особой точки типа «фокус». Внешний устой-

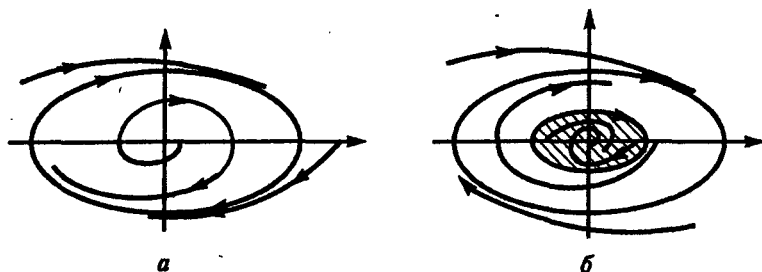


Рис. 11.17. Примеры фазовых портретов с предельным циклом

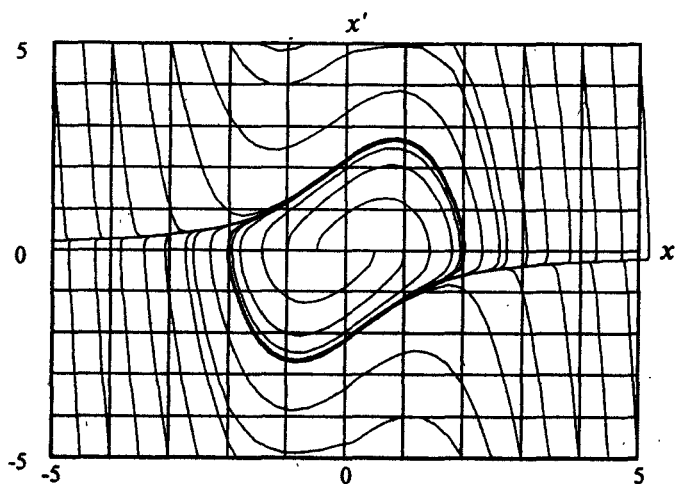


Рис. 11.18. Пример фазового портрета осциллятора Ван дер Поля

чивый предельный цикл разделяет области расходящихся и затухающих движений. При такой топологии фазового пространства говорят о «жестком» режиме возбуждения автоколебаний — для возникновения незатухающих колебаний начальное положение изображающей точки должно быть вне заштрихованной области.

Если фазовый портрет системы имеет более одной особой точки или содержит предельные циклы, то область устойчивости (неустойчивости) не может охватывать всю фазовую плоскость, как это имеет место в линейных системах.

В некоторых случаях границей областей фазовой плоскости с различным характером движений системы служат так называемые *сепаратрисы*. При наличии нескольких особых точек, как это было в примере фазового портрета маятника, сепаратрисами являются траектории, проходящие через седла (см. рис. 11.11).

Фазовые портреты нелинейных систем второго порядка весьма удобны для качественной оценки возможных движений системы. Картина движений в системе определяется топологической структурой фазовой плоскости — типом и взаимным расположением особых траекторий — особых точек, предельных циклов, сепаратрис. Например, по портрету на рис. 11.17, б сразу можно сказать, что положение равновесия устойчиво по отношению к начальным условиям, принадлежащим области притяжения, ограничиваемой неустойчивым предельным циклом. При

начальных условиях за границей области притяжения со временем утанавливаются незатухающие колебания, амплитуда и частота которых локально не зависят от начальных условий. Такая картина сохраняется вне зависимости от размеров предельных циклов; от них зависят только пороговые значения начальных условий, а также амплитуда и частота колебаний, т. е. количественная, а не качественная сторона явлений.

Положим, что построен фазовый портрет некоторой нелинейной системы, который имеет вид, например, показанный на рис. 11.17, б. Если какой-либо параметр q этой системы изменится, то можно построить новый фазовый портрет, который будет либо качественно таким же, либо качественно иным. В последнем случае говорят об изменении топологической структуры фазовой плоскости. Значение параметра q^* , при котором происходит такое изменение, называют *бифуркационным*. Так, например, на фазовом портрете на рис. 11.17, б внутренний предельный цикл может по мере изменения параметра q стягиваться к началу координат и при бифуркационном значении q^* слиться с ним. При дальнейшем изменении q фазовый портрет будет содержать лишь один предельный цикл, а особая точка станет неустойчивой.

Для линейных систем возможны только две топологические структуры фазового пространства: единственная устойчивая или неустойчивая особая точка с неограниченной областью притяжения (отталкивания). Смена этих топологических структур происходит при переходе через критическое значение изменяемого параметра — на границе устойчивости.

Глава 12

УСТОЙЧИВОСТЬ ПОЛОЖЕНИЙ РАВНОВЕСИЯ

При исследовании систем управления по нелинейным математическим моделям следует говорить об устойчивости конкретного движения, а не системы в целом. Действительно, одни движения нелинейной системы могут быть устойчивыми, а другие — нет. Ниже ограничимся рассмотрением устойчивости положений равновесия автономных систем.

12.1. ПОНЯТИЕ ОБ УСТОЙЧИВОСТИ НЕВОЗМУЩЕННОГО ДВИЖЕНИЯ

Пусть математическая модель автономной системы представлена в форме системы дифференциальных уравнений

$$\frac{dv}{dt} = \Phi(v), \quad (12.1)$$

где v — вектор состояния; Φ — вектор-функция. Решения системы (12.1)

$$v^*(t, v_0)$$

при конкретных начальных условиях $v_0 = v(0)$, следуя Ляпунову, называют *невозмущенным движением*. Выбор движения, принимаемого за невозмущенное, произволен. В частности, если начальным условиям v_0 соответствует положение равновесия, то оно и будет невозмущенным движением. Всякое другое решение дифференциальных уравнений

$$v(t, v_0 + \Delta v_0),$$

при иных начальных условиях $v_0 + \Delta v$, называется *возмущенным*.

Вариация движения — разность между возмущенным и невозмущенным движениями:

$$\Delta v(t) = v(t, v_0 + \Delta v_0) - v^*(t, v_0).$$

Невозмущенное движение называется *устойчивым*, если для любого $\varepsilon > 0$ найдется такое $\delta(\varepsilon)$, что из $\|\Delta v_0\| < \delta$ для всех $t > T$ следует: $\|\Delta v(t)\| < \varepsilon$. Здесь $\|\bullet\|$ означает норму вектора.

Если в качестве невозмущенного движения принять состояние равновесия системы, то пользуясь понятием фазового пространства, определе-

ние устойчивости можно интерпретировать следующим образом: положение равновесия называется устойчивым, если, задав вокруг точки равновесия любую сколь угодно малую область ε , можно найти такую область $\delta(\varepsilon)$, что помещенная в момент времени $t = 0$ в любую точку этой области изображающая точка в момент времени $t = T$ войдет в область ε и далее из нее не выйдет. На рис. 12.1 на фазовой плоскости системы второго порядка иллюстрируется приведенное определение для нормы типа

$$\|v\|_{\infty} = \max_i |v_i|.$$

Если принять так называемую евклидову норму

$$\|v\|_2 = \left(\sum_i v_i^2 \right)^{1/2},$$

то области ε и δ в векторных пространствах соответствующих размерностей будут иметь вид шаров с такими радиусами.

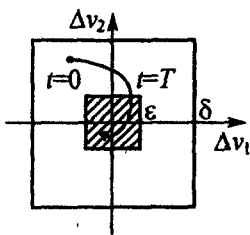


Рис. 12.1. Иллюстрация определения устойчивости

Асимптотическая устойчивость означает, что

$$\lim_{t \rightarrow \infty} \Delta v(t) \rightarrow 0,$$

т. е. движения асимптотически стремятся к невозмущенному, в частности, при $t \rightarrow \infty$ приходят в состояние равновесия.

Устойчивость по Ляпунову — понятие качественное (теоретическое), поскольку не оговариваются размеры области притяжения невозмущенного движения; здесь говорят об устойчивости «в малом». Устойчивость «в большом» — понятие количественное (практическое), когда указываются границы области притяжения. Например, на рис. 11.17, а изображен фазовый портрет системы, где неустойчивый предельный цикл является границей открытой области притяжения устойчивого положения равновесия. В том случае, когда область притяжения совпадает со всем пространством состояний, невозмущенное движение будет устойчивым «в целом».

Линейная система (см. § 3.2) с неособой матрицей A имеет единственное положение равновесия (поэтому здесь говорят об устойчивости системы); если все собственные значения матрицы A левые, то положение равновесия асимптотически устойчиво «в целом». Если все собственные значения левые, кроме некратных на мнимой оси, то положение равновесия устойчиво, но не асимптотически. Разумеется, остается вопрос о том, насколько линейная модель адекватно описывает реакцию системы «в целом».

12.2. ПЕРВЫЙ МЕТОД ЛЯПУНОВА

Метод применяется для исследования устойчивости по линеаризованным уравнениям для малых вариаций переменных.

12.2.1. Применение первого метода Ляпунова к дифференциальным уравнениям в форме Коши

Пусть динамическая система описывается уравнениями (12.1). Обозначим через v^* вектор координат исследуемого положения равновесия, т. е. решение системы уравнений

$$\Phi(v) = 0,$$

и положим, что функции Φ допускают разложения в степенной ряд в точке v^* . Пренебрегая малыми высшего порядка по сравнению с малыми вариациями Δv , получим вместо уравнений (12.1) линеаризованную систему

$$\frac{d\Delta v}{dt} = A(v^*)\Delta v,$$

где

$$A(v^*) = \left. \frac{\partial \Phi(v)}{\partial v} \right|_{v=v^*} = \begin{bmatrix} \frac{\partial \phi_1}{\partial v_1} & \dots & \frac{\partial \phi_1}{\partial v_n} \\ \dots & \dots & \dots \\ \frac{\partial \phi_n}{\partial v_1} & \dots & \frac{\partial \phi_n}{\partial v_n} \end{bmatrix} -$$

— матрица первых производных нелинейных функций (матрица Якоби), вычисляемых в точке равновесия $v = v^*$. Геометрическая иллюстрация процедуры линеаризации в окрестности положения равновесия была приведена на рис. 11.3.

Первый метод Ляпунова базируется на том, что об устойчивости «в малом» положения равновесия нелинейной системы можно судить по результатам анализа линеаризованной системы:

- если все собственные значения матрицы A имеют отрицательные действительные части, т.е. линеаризованная система устойчива асимптотически, то положение равновесия устойчиво;
- если линеаризованная система неустойчива, то положение равновесия неустойчиво.

Первый метод Ляпунова применяется очень часто. Однако он имеет следующие недостатки:

- исследуется только устойчивость «в малом»;
- применим только для систем, линеаризуемых в окрестности положений равновесия.

12.2.2. Исследование устойчивости по дифференциальным уравнениям n -го порядка

Пусть дано дифференциальное уравнение вида (9.4), которое для автономной системы, т.е. при $x(t) \equiv 0$, приводится к виду:

$$\Phi(y, y', \dots, y^{(n)}) = 0.$$

Положения равновесия являются действительными решениями уравнения статики

$$\Phi(y, 0, \dots, 0) = 0,$$

полученного из исходного уравнения приравниванием производных нулю.

Выбирается исследуемое положение равновесия y^* , и левая часть исходного уравнения раскладывается в степенной ряд при условии, что функция Φ аналитична в его окрестности:

$$\Phi(y^*, 0, \dots, 0) + \frac{\partial \Phi}{\partial y} \Big|_* \Delta y + \frac{\partial \Phi}{\partial y'} \Big|_* \Delta y' + \dots + \frac{\partial \Phi}{\partial y^{(n)}} \Big|_* \Delta y^{(n)} + \dots = 0.$$

Полагая, что отклонения переменной y и ее производных малы, можно пренебречь в разложении малыми высших порядков. В результате получится линейное дифференциальное уравнение

$$a_n \frac{d^n \Delta y}{dt^n} + \dots + a_1 \frac{d \Delta y}{dt} + a_0 \Delta y = 0,$$

коэффициенты которого равны значениям частных производных функции Φ в точке равновесия.

Если все корни характеристического полинома

$$A(s) = a_n s^n + \dots + a_1 s + a_0$$

имеют отрицательные действительные части, то положение равновесия исходной нелинейной системы устойчиво.

Пример 1. Проведем исследование устойчивости положения равновесия оспиллятора Ван дер Поля, описываемого дифференциальным уравнением второго порядка:

$$\Phi(y, y', y'') = y'' - \mu(1 - y^2)y' + y = 0.$$

Система имеет единственное положение равновесия $y^* = y' = y'' = 0$. Линеаризованное для малых отклонений уравнение запишется так:

$$a_2 y'' + a_1 y' + a_0 y = 0,$$

где

$$a_0 = \frac{\partial \Phi}{\partial y} \Big|_* = (2\mu y' y + 1) \Big|_* = 1;$$

$$a_1 = \frac{\partial \Phi}{\partial y'} \Big|_* = -\mu(1 - y^2) \Big|_* = -\mu;$$

$$a_2 = \frac{\partial \Phi}{\partial y''} \Big|_* = 1.$$

Характеристический полином уравнения

$$A(s) = s^2 - \mu s + 1$$

имеет следующие корни

$$s_{1,2} = \frac{\mu \pm \sqrt{\mu^2 - 4}}{2}.$$

Положение равновесия не устойчиво, если $\mu > 0$. При значении $\mu \geq 2$ на фазовой плоскости в начале координат имеется особая точка типа «неустойчивый узел». При значениях $0 < \mu < 2$ там же будем иметь особую точку типа «неустойчивый фокус». Фазовый портрет осциллятора Ван дер Поля (см. рис. 11.18) имеет устойчивый предельный цикл, которому соответствуют автоколебания.

В заключение отметим, что в случае нескольких положений равновесия их устойчивость исследуется поочередно. Именно таким был пример математического маятника, рассмотренный в п. 11.3.2.

12.2.3. Исследование устойчивости по моделям в форме структурной схемы

В том случае, когда нелинейная модель представлена в форме структурной схемы, можно, во-первых, записать систему дифференциальных уравнений в форме Коши и исследовать устойчивость положений равновесия по схеме, описанной в § 12.1. Во-вторых, можно попытаться записать одно дифференциальное уравнение n -го порядка; тогда устойчивость исследуется по схеме из п. 12.2.2. Наконец, устойчивость положений равновесия можно исследовать и непосредственно по структурной схеме.

Пусть имеется модель в форме структурной схемы не обязательно расчетного вида; например, она может иметь несколько НЭ и произвольную структуру. Положим, что все НЭ являются безынерционными. Прежде всего, находятся равновесные режимы (если они существуют), устойчивость которых необходимо исследовать. Далее следует рассчитать соответствующие этим состояниям значения переменных на входах всех НЭ системы. Анализ проводится по частной модели для равновесных режимов — динамические звенья заменяются безынерционными. Наличие звеньев интегрирующего или дифференцирующего типа, как было показано в § 10.4, упрощает анализ режимов равновесия. Для определения значений переменных на входах всех НЭ можно применить аналитический, графический или численный методы.

После того как найдено значение $x = x^*$ переменной на входе НЭ, заданного аналитически

$$y = F(x),$$

линеаризация означает его замену безынерционным звеном

$$\Delta y = k \Delta x$$

с коэффициентом усиления

$$k = \left. \frac{dF(x)}{dx} \right|_{x=x^*},$$

равным значению производной в точке равновесия. Ясно, что СХ НЭ в точке линеаризации должна быть гладкой (дифференцируемой).

В том случае, когда СХ НЭ задана графически, линеаризация сводится к определению тангенса угла наклона касательной к СХ в точке равновесия (рис. 12.2). Практически коэффициент усиления определяется как отношение приращений (см. рис. 12.2, а):

$$k = \frac{\Delta y}{\Delta x}.$$

Если НЭ имеет несколько входов

$$y = F(x_1, x_2, \dots, x_N),$$

то в результате линеаризации получим:

$$\Delta y = k_1 \Delta x_1 + k_2 \Delta x_2 + \dots + k_N \Delta x_N,$$

где коэффициенты усиления по различным входам равны значениям частных производных функции F в точке равновесия. Например, линеаризация элемента перемножения двух переменных

$$y = x_1 x_2$$

в точке x_1^*, x_2^* дает

$$\Delta y = x_2^* \Delta x_1 + x_1^* \Delta x_2.$$

В результате линеаризации всех НЭ для конкретного положения равновесия получается линейная модель для малых отклонений перемен-

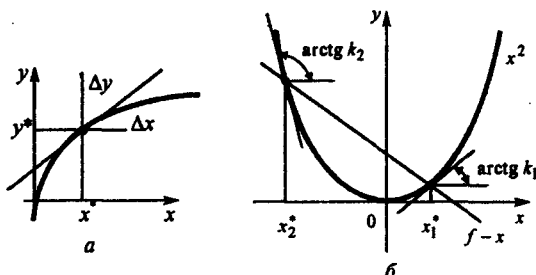


Рис. 12.2. Линеаризация НЭ

ных на входах НЭ. Анализ устойчивости линеаризованной модели в соответствии с первым методом Ляпунова дает возможность судить об устойчивости выбранного положения равновесия нелинейной системы.

Пример 2. Рассмотрим систему (см. рис. 10.10), положения равновесия которой определялись в § 10.5. Пусть передаточная функция ЛЧ имеет вид

$$W(s) = \frac{k}{(s+1)^3},$$

а значение коэффициента усиления $k = 1$. Система имеет два положения равновесия при уровнях постоянного воздействия $f > -0,25$. Значение $f = -0,25$ соответствует катастрофе — малейшее уменьшение f приводит к исчезновению положений равновесия, а малейшее увеличение — к появлению двух положений равновесия. Если $f > -0,25$, то положениям равновесия отвечают следующие значения переменной на входе НЭ:

$$x_{1,2}^* = \frac{-1 \pm \sqrt{1+4f}}{2}.$$

Графическая процедура определения координат $x_{1,2}^*$ точек равновесия, показанная на рис. 12.2, б, соответствует решению системы двух уравнений (10.11) при $k = 1$ или уравнения

$$x^2 = f - x$$

при $f > -0,25$.

Линеаризуем СХ НЭ $y = F(x) \equiv x^2$ в этих точках; получим

$$k_{1,2} = \left. \frac{dF}{dx} \right|_{x=x_{1,2}^*} = 2x|_{x=x_{1,2}^*} = -1 \pm \sqrt{1+4f}.$$

Характеристические полиномы замкнутых линеаризованных систем имеют вид:

$$A_{1,2}(s) = (s+1)^3 + k_{1,2} = s^3 + 3s^2 + 3s + 1 - 1 \pm \sqrt{1+4f} = s^3 + 3s^2 + 3s \pm \sqrt{1+4f}.$$

Видно, что один из полиномов (знак «минус» при свободном члене) не удовлетворяет необходимому условию устойчивости. Для устойчивости другого полинома в соответствии с критерием Гурвица необходимо и достаточно выполнения условия

$$9 > \sqrt{1+4f},$$

т. е. $f < 20$. Таким образом, делаем вывод о том, что из двух положений равновесия, в которых может находиться система при постоянных значениях воздействия $f > -0,25$, устойчивым является одно

$$x_1^* = \frac{-1 + \sqrt{1+4f}}{2}$$

и только при ограниченном сверху уровне воздействия $f < 20$.

12.3. ВТОРОЙ (ПРЯМОЙ) МЕТОД ЛЯПУНОВА

Второй (прямой) метод Ляпунова не связан с линеаризацией дифференциальных уравнений системы.

Пусть автономная система описана дифференциальными уравнениями в форме Коши (12.1). Положим (без потери общности), что положение равновесия $\mathbf{v}^* = 0$; необходимо исследовать его устойчивость.

Введем в рассмотрение однозначную и дифференцируемую в некоторой области, содержащей начало координат, скалярную функцию переменных состояния

$$V(\mathbf{v}) = V(v_1, \dots, v_n).$$

Функция $V(\mathbf{v})$ называется *знакопостоянной*, если она имеет один и тот же знак всюду в некоторой области, содержащей начало координат ($\mathbf{v} = 0$), за исключением некоторых точек, где она равна нулю. Знакопостоянная функция, равная нулю только в начале координат, называется *знакоопределенной* (положительно-определенной или отрицательно-определенной — в зависимости от знака). На рис. 12.3 показан пример положительно-определенной функции

$$V(v_1, v_2) = v_1^2 + v_2^2.$$

А. М. Ляпунов показал, что для устойчивости положения равновесия *достаточно* существования положительно-определенной функции $V(\mathbf{v})$, полная производная по времени которой $dV(\mathbf{v})/dt$ в силу дифференциальных уравнений системы является отрицательной *знакопостоянной* функцией $W(\mathbf{v})$.

Если найдена положительно-определенная функция, для которой выполняется сформулированное достаточное условие, то такую функцию называют функцией Ляпунова.

Если функция $W(\mathbf{v})$ отрицательно-определенная, то положение равновесия устойчиво асимптотически. Если это условие выполнено при любых $\mathbf{v}$ и, кроме того, $V(\mathbf{v}) \rightarrow \infty$ при $\|\mathbf{v}\| \rightarrow \infty$, то положение равновесия асимптотически устойчиво в целом.

Не будем приводить доказательств этих утверждений, дадим только некоторые пояснения.

Производная по времени положительно-определенной функции

$$\frac{dV(\mathbf{v})}{dt} = \frac{\partial V(\mathbf{v})}{\partial \mathbf{v}} \cdot \frac{d\mathbf{v}}{dt}$$

в силу дифференциальных уравнений (12.1) равна:

$$\frac{dV(\mathbf{v})}{dt} = \frac{\partial V(\mathbf{v})}{\partial \mathbf{v}} \Phi(\mathbf{v}) = W(\mathbf{v}),$$

т. е. представляет собой функцию только переменных состояния системы. Свойства этой функции и определяют устойчивость положения равновесия. Функция $W(\mathbf{v})$ равна скалярному произведению вектора градиента функции $V(\mathbf{v})$

$$\text{grad} V(\mathbf{v}) = \left[\frac{\partial V(\mathbf{v})}{\partial \mathbf{v}} \right]$$

на вектор скорости изменения состояния системы $d\mathbf{v}/dt$. Известно, что скалярное произведение векторов равно произведению их модулей, умноженному на косинус угла между векторами:

$$W(\mathbf{v}) = |\text{grad} V(\mathbf{v})| \cdot |d\mathbf{v}/dt| \cos \psi.$$

Если по условиям утверждений функция $W(\mathbf{v})$ — отрицательная знакопостоянная (знакоопределенная), то это значит, что угол ψ неострый (тупой). Другими словами, если проекция вектора скорости изменения состояния системы на вектор градиента функции $V(\mathbf{v})$ неположительна (отрицательна), то состояния системы меняются только в направлении невозрастания (убывания) функции Ляпунова.

На рис. 12.4 показана фазовая плоскость некоторой системы второго порядка ($n = 2$) с нанесенными на ней линиями равного уровня функции

$$V(v_1, v_2) = c_i = \text{const.}$$

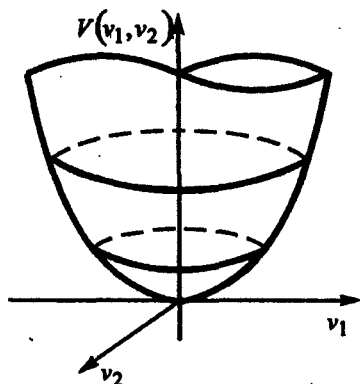


Рис. 12.3. Пример положительно-определенной функции

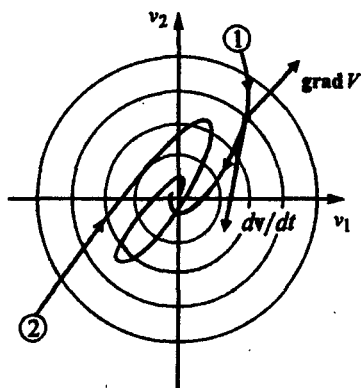


Рис. 12.4. Иллюстрация второго метода Ляпунова

Это вложенные друг в друга всюду плотно замкнутые линии, с уменьшением c , стягивающиеся к началу координат. Для выбранной точки фазовой плоскости на рисунке показан вектор $\text{grad} V(\mathbf{v})$. Через эту точку проходит единственная фазовая траектория (1), касательная к которой есть вектор скорости $(dv_1/dt, dv_2/dt)^T$. Если угол между этими векторами тупой, то фазовая траектория системы переходит от больших значений V к меньшим и с течением времени стремится к началу координат (точке равновесия).

Значение функции $V(\mathbf{v})$ можно интерпретировать как некоторое обобщенное расстояние между состоянием системы и положением равновесия, устойчивость которого исследуется. Уменьшение этого расстояния вдоль любой фазовой траектории гарантирует устойчивость.

Изложенный метод базируется на достаточном условии устойчивости, т. е. его невыполнение не означает, что невозмущенное движение неустойчиво. Сказанное иллюстрируется на рис. 12.4 фазовой траекторией (2), для которой условия второго метода при выбранной функции $V(\mathbf{v})$ не выполняются, хотя отклонения при конкретных начальных условиях затухают, т. е., быть может, положение равновесия устойчиво.

Пример 3. Проведем исследование устойчивости положения равновесия системы уравнений осциллятора с нелинейным демпфированием:

$$\begin{aligned}\frac{dv_1}{dt} &= v_2; \\ \frac{dv_2}{dt} &= -v_1 - v_2^3.\end{aligned}$$

Попытаемся исследовать устойчивость положения равновесия ($v_1 = v_2 = 0$) с помощью первого метода Ляпунова. Линеаризация дает систему уравнений для малых отклонений, которую представим в векторно-матричной форме, приняв вектор состояния $\mathbf{v} = (v_1, v_2)^T$:

$$\frac{d\mathbf{v}}{dt} = \mathbf{A}\mathbf{v} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{v}.$$

Матрица системы имеет мнимые собственные значения $\pm j$, т. е. линеаризованная система находится на границе устойчивости. В этом критическом случае об устойчивости положения равновесия исходной нелинейной системы нельзя судить по линеаризованным уравнениям (см. п. 12.2.1).

Воспользуемся вторым методом Ляпунова. Выберем скалярную функцию переменных состояния в виде квадратичной формы:

$$V(\mathbf{v}) = v_1^2 + v_2^2$$

и продифференцируем по времени в силу дифференциальных уравнений системы:

$$\frac{dV(\mathbf{v})}{dt} = 2v_1v_2 + 2v_2(-v_1 - v_2^3) = -2v_2^4 \leq 0.$$

Получили отрицательно-определенную функцию, что означает асимптотическую устойчивость положения равновесия. Поскольку при $\|v\| \rightarrow \infty$ имеем $V(v) \rightarrow \infty$, положение равновесия устойчиво в целом.

Отметим, что иллюстрировать асимптотическую сходимость процессов к нулю имитационным способом также не просто — с приближением к началу координат относительное затухание процессов уменьшается, так как ослабляется влияние нелинейного демпфирующего члена v^3 .

12.4. ПРИМЕНЕНИЕ ВТОРОГО МЕТОДА ЛЯПУНОВА

Второй метод Ляпунова универсален, так как не связан с линеаризацией уравнений движения (12.4) и не накладывает особых ограничений на их правые части. Вместе с тем, применение второго метода Ляпунова в практике проектирования систем управления осложняется двумя обстоятельствами:

1) достаточным характером утверждений (т. е. если условия метода не выполняются, то об устойчивости положения равновесия ничего сказать нельзя, можно только порекомендовать подобрать другую функцию $V(v)$;

2) отсутствием общих рекомендаций по выбору функций Ляпунова.

Чаще всего в качестве функции $V(v)$ выбирают квадратичные формы

$$V(v) = v^T H v, \quad (12.2)$$

где H — положительно-определенная матрица. Выражение (12.2) для $n = 2$ раскрывается так:

$$V(v_1, v_2) = (v_1 v_2) \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = h_{11}v_1^2 + h_{12}v_1v_2 + h_{21}v_2v_1 + h_{22}v_2^2.$$

Для установления положительной определенности матрицы H можно воспользоваться критерием Сильвестра, сводящимся к проверке положительности диагональных определителей матрицы. Например, для $n = 2$ условия запишутся так:

$$h_{11} > 0; \quad h_{11}h_{22} - h_{12}h_{21} > 0; \quad h_{22} > 0.$$

Недостатком функции (12.2) является то, что она не учитывает особенностей нелинейных систем.

Рассмотрим нелинейную систему расчетного вида (см. рис. 9.13). Пусть СХ $F(x)$ безынерционного НЭ удовлетворяет следующим условиям:

- она однозначна и непрерывна;
- $F(0) = 0$;
- $xF(x) > 0, x \neq 0$,

т. е. график СХ проходит через начало координат и располагается в первом и третьем квадрантах. Для этого практически важного случая А. И. Лурье и В. Н. Постников предложили следующую форму функции Ляпунова (квадратичная форма плюс интеграл от нелинейности):

$$V(v) = v^T H v + \beta \int_0^x F(\xi) d\xi.$$

Пример 4. Пусть ЛЧ системы (см. рис. 9.13) имеет передаточную функцию

$$W(s) = k_0 / (T_0 s + 1),$$

а НЭ удовлетворяет приведенным условиям. При отсутствии воздействия ($f = 0$) положению равновесия системы соответствует $x = 0$.

Дифференциальное уравнение системы первого порядка в форме Коши запишется так (примем $v = x$):

$$\frac{dx}{dt} = -\frac{1}{T_0} (x + k_0 F(x)).$$

Выберем следующую функцию Ляпунова:

$$V(x) = x \frac{1}{2} x + k_0 \int_0^x F(\xi) d\xi. \quad (12.3)$$

Продифференцируем эту функцию по времени в силу дифференциального уравнения системы:

$$\frac{dV(x)}{dx} \cdot \frac{dx}{dt} = -\frac{1}{T_0} [x + k_0 F(x)]^2 = W(x).$$

Получили отрицательно-определенную функцию $W(x)$, что позволяет сделать вывод об асимптотической устойчивости положения равновесия. Кроме того, замечаем, что функция (12.3) определена для всех x и при $|x| \rightarrow \infty$ имеем $V(x) \rightarrow \infty$. Поэтому положение равновесия асимптотически устойчиво в целом. Наконец, примем во внимание, что полученный результат справедлив для целого класса нелинейных функций $F(x)$, удовлетворяющих введенным выше ограничениям.

Фазовый портрет рассмотренной системы первого порядка представляет собой единственную кривую

$$x' = -(x + k_0 F(x)) / T_0,$$

проходящую через начало координат и располагающуюся во втором и четвертом квадрантах (при условии $k_0 > 0$, $T_0 > 0$). Во втором квадранте ($x < 0$, $x' > 0$) изображающая точка движется в сторону увеличения значений x , а в четвертом ($x > 0$, $x' < 0$) — в сторону уменьшения значений x , т. е. при любых начальных условиях изображающая точка движется к началу координат — положение равновесия устойчиво асимптотически в целом.

Асимптотическую устойчивость в целом для класса нелинейностей называют *абсолютной устойчивостью*. В рассмотренном примере 4 системы первого порядка положение равновесия оказалось абсолютно устойчивым.

12.5. ЧАСТОТНЫЙ МЕТОД ИССЛЕДОВАНИЯ АБСОЛЮТНОЙ УСТОЙЧИВОСТИ

Рассмотрим нелинейную систему расчетной структуры (см. рис. 9.13) с устойчивой ЛЧ. Положим, что НЭ имеет однозначную СХ, удовлетворяющую условиям:

$$\begin{aligned} F(0) &= 0; \\ 0 \leq F(x) &\leq kx. \end{aligned} \quad (12.4)$$

Это означает принадлежность СХ секторам в первом и третьем квадрантах (рис. 12.5), образованным осью абсцисс и прямой $y = kx$. Говорят, что НЭ принадлежит углу $[0, k]$. Это является основным признаком рассматриваемого класса нелинейностей.

При отсутствии внешнего воздействия система имеет единственное положение равновесия $x = 0$, устойчивость которого исследуется далее.

Необходимое условие абсолютной устойчивости. Известна выдвинутая М. А. Айзерманом гипотеза о том, что если заменить НЭ линейным усилительным звеном $y = k_{\text{н}}x$, причем, линеаризованная таким образом система устойчива при всех $0 \leq k_{\text{н}} \leq k$ (как говорят, k — гурвицев угол), то положение равновесия системы с любым НЭ, принадлежащим углу $[0, k]$ будет абсолютно устойчивым. Однако были найдены примеры (В. А. Плисс), когда положения равновесия систем, для которых выполняются условия гипотезы, не были устойчивы в целом.

Очевидно, условия гипотезы являются *необходимыми*, поскольку линейные преобразования являются частным случаем нелинейных и принадлежат описанному выше классу (12.4). Необходимое условие абсолютной устойчивости легко проверить. Из теории линейных систем управления (см. § 3.5) известно, что рассматриваемая система устойчива при любом значении усиления $0 \leq k_{\text{н}} \leq k$, если АФХ ЛЧ не пересекает луча $(-\infty, -1/k]$ на действительной отрицательной оси.

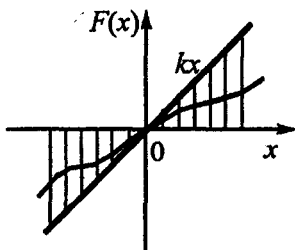


Рис. 12.5. Нелинейность в угле

Достаточное условие абсолютной устойчивости. Частотный критерий абсолютной устойчивости положения равновесия впервые был сформулирован В.-М. Поповым (V.-M. Popov): для абсолютной устойчивости положения равновесия нелинейной системы с устойчивой ЛЧ *достаточно* существования действительного q , для которого выполняется условие

$$\forall \omega \geq 0: \operatorname{Re}[(1 + j\omega q)W(j\omega)] > -1/k, \quad (12.5)$$

где k — угол абсолютной устойчивости. Формулировка критерия (12.5) допускает простую геометрическую интерпретацию. Представим АФХ ЛЧ в виде:

$$W(j\omega) = P(\omega) + jQ(\omega). \quad (12.6)$$

После подстановки (12.6) в неравенство (12.5) и выделения действительной части получим условие абсолютной устойчивости в виде:

$$P(\omega) - q\omega Q(\omega) > -1/k. \quad (12.7)$$

Вводится понятие *модифицированной АФХ* ЛЧ

$$W^*(\omega) = P^*(\omega) + jQ^*(\omega),$$

где

$$P^*(\omega) = P(\omega), \quad Q^*(\omega) = \omega Q(\omega). \quad (12.8)$$

Модифицированная АФХ отличается от обычной изменением значений мнимой части в ω раз.

Подставив (12.8) в условие (12.7), получим достаточное условие абсолютной устойчивости для модифицированной АФХ ЛЧ системы:

$$P^*(\omega) - qQ^*(\omega) > -1/k. \quad (12.9)$$

Введем в рассмотрение так называемую прямую Попова на плоскости модифицированной АФХ:

$$P^*(\omega) - qQ^* = -1/k$$

или

$$Q^* = (P^* + 1/k)/q.$$

Прямая проходит через точку $-1/k$ на действительной оси с наклоном $1/q$. Когда выполняется условие (12.9), вся модифицированная АФХ $W^*(\omega)$ лежит правее прямой Попова.

Пример 5. Пусть ЛЧ системы описывается типовым аperiодическим звеном

$$W(s) = \frac{k_0}{Ts + 1}.$$

Частотная характеристика ЛЧ определяется выражением

$$W(j\omega) = P(\omega) + jQ(\omega) = \frac{k_0}{1 + \omega^2 T^2} - j \frac{k_0 \omega T}{1 + \omega^2 T^2}.$$

Мнимая часть модифицированной АФХ равна

$$Q^*(\omega) = \omega Q(\omega) = -\frac{k_0 \omega^2 T}{1 + \omega^2 T^2}.$$

Исключая ω из выражения для $P^*(\omega)$ и Q^* , получим уравнение модифицированной АФХ в виде отрезка прямой (рис. 12.6):

$$Q^* = (P^* - k_0)/T.$$

Прямую Попова при любом положительном значении q можно провести даже через начало координат, что соответствует $k \rightarrow \infty$. Таким образом, положение равновесия абсолютно устойчиво для всех НЭ с однозначной СХ, лежащей в угле $[0, \infty)$ (т. е. в первом и третьем квадрантах). Такой же вывод был сделан ранее в результате решения этой задачи вторым методом Ляпунова.

Частотный критерий абсолютной устойчивости применим и к нелинейным системам, НЭ которых имеют неоднозначные СХ [88]. Для случая НЭ с гистерезисом действительное число q в условии (12.5) при этом должно быть неотрицательным ($q \geq 0$), а в случае НЭ с опережением в условии (12.5) принимается $q \leq 0$.

Обратим внимание на достаточный характер критерия абсолютной устойчивости в частотной форме. При исследовании нелинейных систем могут быть три ситуации. Во-первых, не выполняется необходимое условие абсолютной устойчивости (рис. 12.7, а) — обычная АФХ ЛЧ пересекает луч $(-\infty, -1/k)$. Во-вторых, модифицированная АФХ лежит правее прямой Попова (рис. 12.7, б) — положение равновесия абсолютно устойчиво. В-третьих, выполняется необходимое условие, но не выполняется достаточное (рис. 12.7, в) — нельзя провести прямую Попова. Здесь об устойчивости положения равновесия ничего сказать нельзя.

Круговой критерий. Пусть НЭ имеет нестационарную характеристику

$$y = F(x, t),$$

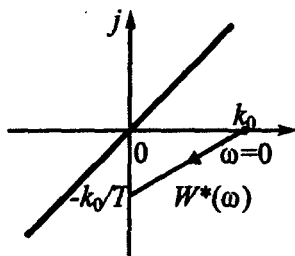


Рис. 12.6. Пример исследования абсолютной устойчивости

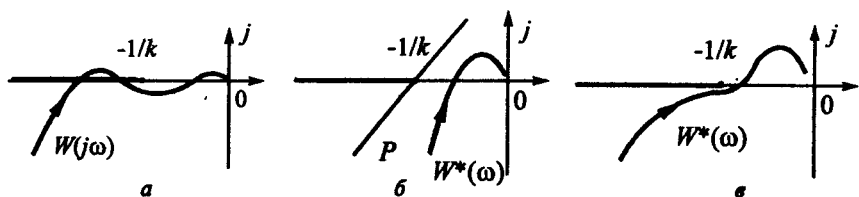


Рис. 12.7. Ситуации исследования абсолютной устойчивости

принадлежащую сектору $[r, k]$ для всех t (рис. 12.8), т. е. удовлетворяющую условиям:

$$F(0, t) = 0; \quad rx \leq F(x, t) \leq kx.$$

Положение равновесия нелинейной системы с нестационарным НЭ абсолютно устойчиво, если АФХ устойчивой ЛЧ не охватывает точек круга с центром на действительной оси, показанного на рис. 12.9.

Следует обратить внимание на то, что круговой критерий формулируется для обычных (не модифицированных) АФХ ЛЧ. Обратим также внимание на то, что при $r \rightarrow k$, когда СХ НЭ «выпрямляется», круговой критерий совпадает с критерием Найквиста. В том же случае, когда $r = 0$, т. е. СХ НЭ принадлежит углу $[0, k]$, круг превращается в полуплоскость, а условие абсолютной устойчивости — в требование

$$\operatorname{Re} W(j\omega) > -1/k.$$

Частотный метод исследования абсолютной устойчивости весьма удобен не только для анализа, но и для синтеза нелинейных систем. В частности, при заданной ЛЧ легко определить максимальный угол k сектора, которому должны принадлежать СХ НЭ систем с абсолютно устойчивым положением равновесия.

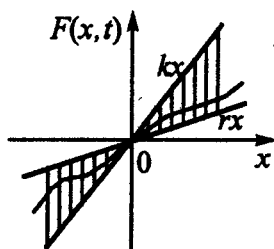


Рис. 12.8. Нестационарные НЭ в угле

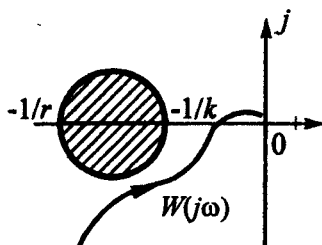


Рис. 12.9. Иллюстрация кругового критерия

ИССЛЕДОВАНИЕ ПЕРИОДИЧЕСКИХ РЕЖИМОВ МЕТОДОМ ГАРМОНИЧЕСКОГО БАЛАНСА

13.1. ОСНОВНЫЕ ПОЛОЖЕНИЯ МЕТОДА ГАРМОНИЧЕСКОГО БАЛАНСА

Периодические процессы в нелинейных системах — автоколебания — являются часто встречающимися и практически важными режимами их функционирования. Если система, рассчитанная по линейным моделям, теряет устойчивость, то, как правило, колебания расходятся до определенной амплитуды (см. рис. 9.2, б). Автоколебательные системы успешно применяются для поддержания различных физических переменных технологических процессов (например, температуры), если амплитуда и частота колебаний находятся в допустимых пределах.

Точное определение формы и параметров периодических режимов возможно только для некоторых типов нелинейных систем, в частности релейных. Для исследования предельных циклов в системах второго порядка весьма удобен метод фазовой плоскости. Однако во многих случаях системы управления описываются моделями высоких порядков и имеют сложные нелинейные характеристики. В этих случаях применяют приближенные методы исследования периодических режимов [19, 52, 60, 72, 86].

Важную информацию о существовании периодических режимов в нелинейных системах, их числе и параметрах может дать приближенный *метод гармонического баланса*. Достоинство частотного метода гармонического баланса заключается в его наглядности, физичности, в возможности получения зависимости показателей качества процессов от вида и параметров нелинейности, структуры и параметров линейной части, что является предпосылкой решения задач синтеза. По результатам, полученным этим методом, могут быть оценены начальные условия для моделирования систем на ЭВМ с целью последующего уточнения форм и параметров локализованных периодических режимов.

Пусть имеется нелинейная модель автономной системы, представленная структурной схемой расчетного вида (см. рис. 9.13). Поставлена задача исследования собственных периодических колебаний.

Для упрощения задачи искомое решение параметризуют, т. е. принимают гипотезу о его форме. Предположим, что переменная на входе $x(t)$ НЭ периодична, а ее форма близка к гармонической:

$$x(t) = A \sin \omega t. \quad (13.1)$$

Проанализируем условия прохождения гармонического сигнала по контуру, образованному НЭ, ЛЧ и элементом сравнения (см. рис. 9.13). Допустим, что периодический сигнал на выходе НЭ

$$y(t) = F(x(t))$$

имеет тот же период. Однако он содержит высшие гармоники, которые, проходя через ЛЧ с передаточной функцией $W(s)$, ослабляются в большей степени, чем первая гармоническая составляющая (гипотеза фильтра). Поэтому в составе переменной $z(t)$ их доля мала:

$$z(t) \approx A_z \sin(\omega t + \varphi_z).$$

С учетом отрицательной обратной связи, изменяющей фазу сигнала на $-\pi$ радиан, условием существования установившихся колебаний в контуре является баланс амплитуд и фаз:

$$\begin{aligned} A_z &= A; \\ \varphi_z &= -\pi. \end{aligned} \quad (13.2)$$

Комплексный коэффициент передачи ЛЧ по гармонической составляющей выражается так:

$$W(j\omega) = R(\omega) \exp(j\varphi(\omega)). \quad (13.3)$$

По аналогии примем, что преобразование гармонического сигнала безынерционным НЭ описывается комплексным коэффициентом передачи (описывающей функцией)

$$W_n(A) = R_n(A) \exp(j\varphi_n(A)), \quad (13.4)$$

где $R_n(A)$ — отношение амплитуды первой гармонической составляющей сигнала $y(t)$ к амплитуде A входного сигнала; $\varphi_n(A)$ — фазовый сдвиг первой гармоники по отношению к входному сигналу (13.1). Таким образом, при прохождении гармонического сигнала через последовательно соединенные НЭ и ЛЧ его амплитуда изменяется в $R(\omega)R_n(A)$ раз, а сдвиг фаз равен $\varphi(\omega) + \varphi_n(A)$.

В нелинейной системе (см. рис. 9.13) возможно существование периодических колебаний, если при прохождении гармонического сигнала вдоль всего контура амплитуда сохраняет свое значение, а фазовый сдвиг с учетом элемента сравнения равен $-\pi$, т. е. выполняются условия

$$\begin{aligned} R_n(A)R(\omega) &= 1, \\ \varphi_n(A) + \varphi(\omega) &= -\pi. \end{aligned} \quad (13.5)$$

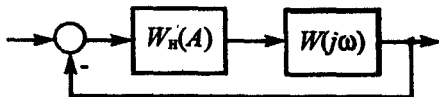


Рис. 13.1. Эквивалентная система

К получению условий (13.5) можно подойти и несколько иначе. На рис. 13.1 показана структурная схема гармонически линеаризованной системы, эквивалентной исходной только для множества гармонических сигналов (13.1) на входе НЭ и при выполнении гипотезы фильтра. Особенностью этой частной модели является неполная определенность ее параметров, зависящих от искомой амплитуды A и частоты ω . В исходной системе возможны периодические колебания, если в эквивалентной системе, показанной на рис. 13.1, устанавливаются гармонические колебания, т. е. параметры линеаризованной системы соответствуют колебательной границе устойчивости. В соответствии с критерием Найквиста условие нахождения линеаризованной системы, заданной частотными характеристиками звеньев, на границе устойчивости запишется следующим образом:

$$W_n(A)W(j\omega) = -1, \quad (13.6)$$

из которого с учетом (13.3) и (13.4) вытекают условия (13.5) баланса амплитуд и фаз.

Условия (13.5) представляют собой систему нелинейных уравнений относительно частоты ω и амплитуды A искомой первой гармоники периодических колебаний переменной $x(t)$.

Принятая гипотеза фильтра — гипотеза о гармонической форме искомого периодического движения — говорит о малости амплитуд высших гармоник на входе НЭ и выходе ЛЧ. Другое необходимое условие применимости метода — малость влияния высших гармоник и неучтенных параметров системы на условие баланса амплитуд и фаз [19].

В общем случае баланс может устанавливаться с учетом постоянного смещения x_0 гармонического сигнала на входе НЭ и конечного числа высших гармоник. Система соответствующих нелинейных уравнений с большим числом неизвестных составляется и решается с помощью ЭВМ.

В случае принятия гипотезы о близости формы искомым колебаний к гармонической основными этапами метода гармонического баланса являются:

- замена НЭ эквивалентной характеристикой (гармоническая линеаризация);
- определение параметров колебаний;

- исследование устойчивости колебаний;
- анализ чувствительности решения к высшим гармоникам и малым параметрам системы.

13.2. ГАРМОНИЧЕСКАЯ ЛИНЕАРИЗАЦИЯ НЕЛИНЕЙНОГО ЭЛЕМЕНТА

По существу, речь идет об эквивалентировании НЭ для множества гармонических сигналов. Будем рассматривать случай симметричных колебаний (13.1).

Коэффициенты гармонической линеаризации определяются как отношение первых гармонических составляющих выхода НЭ к амплитуде входа:

$$\begin{aligned} q(A, \omega) &= \frac{\alpha}{A} = \frac{2}{TA} \int_0^T F(A \sin \omega t) \sin \omega t dt, \\ q'(A, \omega) &= \frac{\beta}{A} = \frac{2}{TA} \int_0^T F(A \sin \omega t) \cos \omega t dt, \end{aligned} \quad (13.7)$$

где α и β — коэффициенты разложения Фурье по синусной и косинусной составляющим. Формулы (13.7) означают, что гармоническая линеаризация дает усредненные за период свойства НЭ по преобразованию гармонического сигнала.

Получим выражение для эквивалентной комплексной характеристики НЭ по гармоническому сигналу. Первая гармоническая составляющая периодического сигнала на выходе НЭ равна

$$y_1(t) = [q(A) \sin \omega t + q'(A) \cos \omega t] A.$$

Учтем, что

$$\cos \omega t = \frac{1}{\omega} \cdot \frac{d}{dt} \sin \omega t = \frac{1}{\omega} p \sin \omega t,$$

где $p \equiv d/dt$ — оператор дифференцирования по времени, и запишем с учетом (13.1):

$$y_1(t) = \left[q(A) + \frac{p}{\omega} q'(A) \right] x(t).$$

В квадратных скобках — эквивалентный оператор НЭ; можно также говорить об эквивалентной передаточной функции

$$W_H(A, s) = q(A) + \frac{s}{\omega} q'(A) \quad (13.8)$$

и эквивалентной частотной характеристике ($s = j\omega$):

$$W_H(A) = q(A) + jq'(A) = R_H(A) \exp(j\varphi_H(A)),$$

модуль и аргумент которой вычисляются так:

$$R_H(A) = \sqrt{q^2(A) + [q'(A)]^2}; \quad (13.9)$$

$$\varphi_H(A) = \text{Arctg}(q'(A)/q(A)). \quad (13.10)$$

13.3. СПОСОБЫ ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТОВ ГАРМОНИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

Для динамических НЭ установившийся периодический сигнал при гармоническом входе (13.1) с конкретными значениями A и ω определяется путем численного решения дифференциальных уравнений, после чего по формулам (13.7) находят значения $q(A, \omega)$, $q'(A, \omega)$. Повторяя эту операцию, можно получить таблицу значений коэффициентов гармонической линеаризации.

Возможно также вычисление коэффициентов гармонической линеаризации безынерционных НЭ усреднением по фазе, а также по множеству значений входа.

13.3.1. Вычисление коэффициентов гармонической линеаризации усреднением по фазе

В частном случае безынерционных НЭ коэффициенты гармонической линеаризации от частоты ω не зависят. Поэтому от усреднения по времени (13.7) можно перейти к усреднению по фазе:

$$\begin{aligned} q(A) &= \frac{1}{\pi A} \int_0^{2\pi} F(A \sin \psi) \sin \psi d\psi; \\ q'(A) &= \frac{1}{\pi A} \int_0^{2\pi} F(A \sin \psi) \cos \psi d\psi. \end{aligned} \quad (13.11)$$

Практическое применение формул (13.11) сводится к получению периодического сигнала

$$y(\psi) = F(A \sin \psi)$$

и вычислению интегралов. Для НЭ с простыми СХ сигнал $y(\psi)$ строится графически, а интегралы (13.11) выражаются через элементарные функции.

Пример 1. Получим выражения для коэффициентов гармонической линейризации идеального релейного элемента (см. рис. 9.6). Периодический сигнал $y(\psi)$ при гармоническом входе (13.1) представляет собой прямоугольные колебания с уровнями $\pm C$.

$$y(\psi) = \begin{cases} C, & 0 \leq \psi < \pi; \\ -C, & \pi \leq \psi < 2\pi. \end{cases}$$

По формулам (13.11) получим:

$$q(A) = \frac{1}{\pi A} \left(\int_0^{\pi} C \sin \psi d\psi - \int_{\pi}^{2\pi} C \sin \psi d\psi \right) = \frac{C}{\pi A} \left(-\cos \psi \Big|_0^{\pi} + \cos \psi \Big|_{\pi}^{2\pi} \right) = \frac{4C}{\pi A}.$$

$$q'(A) = \frac{1}{\pi A} \left(\int_0^{\pi} C \cos \psi d\psi - \int_{\pi}^{2\pi} C \cos \psi d\psi \right) = \frac{C}{\pi A} \left(\sin \psi \Big|_0^{\pi} - \sin \psi \Big|_{\pi}^{2\pi} \right) = 0.$$

Эквивалентная амплитудная характеристика идеального реле в силу (13.9) равна $q(A)$:

$$R_n(A) = \frac{4C}{\pi A}, \quad (13.12)$$

а фазовые сдвиги $\varphi_n(A)$ равны нулю, т.е. фаза первой гармонической составляющей выхода совпадает с фазой входа.

График функции (13.12) показан на рис. 13.2, из которого видно, что малые по амплитуде сигналы проходят с большим усилением, а большие — с малым.

Особенность безынерционного нелинейного преобразования — дискриминация сигналов по уровню.

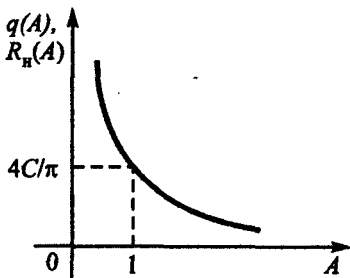


Рис. 13.2. График зависимости эквивалентного усиления идеального реле от амплитуды гармонического сигнала

13.3.2. Вычисление коэффициентов гармонической линейризации усреднением по множеству значений входа нелинейного элемента

В случае безынерционных НЭ для вычисления коэффициентов гармонической линейризации не обязательно получение периодического сигнала на выходе. Пусть НЭ имеет двузначную СХ (рис. 13.3) с восходящей $F_1(x)$ и нисходящей $F_2(x)$ ветвями.

Проведем замену переменных в интегралах (13.11):

$$A \sin \psi = x;$$

$$\sin \psi = x/A; \quad dx = A \cos \psi d\psi; \quad \cos \psi d\psi = dx/A;$$

$$\psi = \text{Arcsin}(x/A);$$

$$d\psi = \pm dx / \sqrt{A^2 - x^2},$$

причем знак «плюс» имеет место для значений ψ в интервалах $(0, \pi/2)$; $(3\pi/2, 2\pi)$ а знак «минус» — в интервалах $(\pi/2, \pi)$; $(\pi, 3\pi/2)$. Пределы интегрирования $(0, 2\pi)$ разобьем по соответствующим квадрантам: $0; \pi/2; \pi; 3\pi/2; 2\pi$. Этим интервалам (с учетом замены переменной) соответствуют следующие пределы интегрирования по x :

$$(0, A); (A, 0); (0, -A); (-A, 0).$$

В результате формулы (13.11) для коэффициентов гармонической линеаризации запишутся так:

$$\begin{aligned} q(A) &= \frac{1}{\pi A^2} \left[\int_0^A \frac{F_1(x) x dx}{\sqrt{A^2 - x^2}} - \int_A^0 \frac{F_2(x) x dx}{\sqrt{A^2 - x^2}} - \int_0^{-A} \frac{F_2(x) x dx}{\sqrt{A^2 - x^2}} + \int_{-A}^0 \frac{F_1(x) x dx}{\sqrt{A^2 - x^2}} \right] = \\ &= \frac{1}{\pi A^2} \left[\int_{-A}^A \frac{F_1(x) x dx}{\sqrt{A^2 - x^2}} + \int_{-A}^A \frac{F_2(x) x dx}{\sqrt{A^2 - x^2}} \right] = \frac{1}{\pi A^2} \int_{-A}^A \frac{[F_1(x) + F_2(x)]}{\sqrt{A^2 - x^2}} x dx; \end{aligned} \quad (13.13)$$

$$\begin{aligned} q'(A) &= \frac{1}{\pi A^2} \left[\int_0^A F_1(x) dx + \int_A^0 F_2(x) dx + \int_0^{-A} F_2(x) dx + \int_{-A}^0 F_1(x) dx \right] = \\ &= \frac{1}{\pi A^2} \left[\int_{-A}^A F_1(x) dx + \int_{-A}^A F_2(x) dx \right] = \frac{1}{\pi A^2} \int_{-A}^A [F_1(x) + F_2(x)] dx. \end{aligned} \quad (13.14)$$

Формулы (13.13) и (13.14) справедливы, если амплитуда A входного сигнала охватывает весь интервал двузначности СХ НЭ. Эти формулы, по существу, означают усреднение по множеству значений x сигнала на входе НЭ.

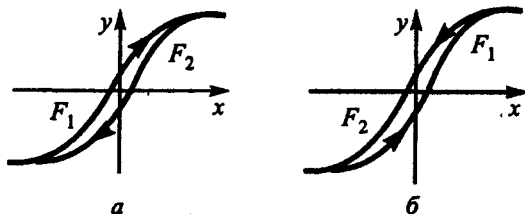


Рис. 13.3. Нелинейности с двузначными СХ

Обратим внимание на то, что коэффициент $q'(A)$ оказывается пропорциональным площади между ветвями СХ. Пусть знак $q(A)$ всегда положителен. Тогда из условия $F_1(x) \geq F_2(x)$ для всех x следует $q'(A) \geq 0$ и наоборот. Это значит, что по виду СХ можно судить о знаках $q'(A)$ и фазового сдвига: НЭ, в которых восходящая ветвь СХ проходит не ниже нисходящей, вносят по первой гармонической составляющей положительный фазовый сдвиг (рис. 13.3, а), а НЭ, в которых восходящая ветвь СХ проходит ниже нисходящей, вносят отрицательный фазовый сдвиг (рис. 13.3, б).

Если СХ НЭ однозначна, т.е. $F_1(x) \equiv F_2(x) \equiv F(x)$, то из (13.13) и (13.14) получаем:

$$q(A) = \frac{2}{\pi A^2} \int_{-A}^A \frac{F(x)xdx}{\sqrt{A^2 - x^2}};$$

$$q'(A) \equiv 0.$$

Как следует из формул (13.9), (13.10), для однозначных НЭ справедливы соотношения:

$$R_n(A) = q(A);$$

$$\Phi_n(A) \equiv 0.$$

Если НЭ имеет однозначную нечетную СХ, т.е. $F(-x) = -F(x)$, то выражение для $q(A)$ принимает вид:

$$q(A) = \frac{4}{\pi A^2} \int_0^A \frac{F(x)xdx}{\sqrt{A^2 - x^2}} = R_n(A). \quad (13.15)$$

Пример 2. Получим выражения для коэффициентов гармонической линеаризации идеального реле. Поскольку СХ — однозначная нечетная, то $q'(A) \equiv 0$; для получения $q(A)$ воспользуемся формулой (13.15):

$$q(A) = \frac{4}{\pi A^2} \int_0^A \frac{Cx dx}{\sqrt{A^2 - x^2}} = \frac{4C}{\pi A^2} \left(-\sqrt{A^2 - x^2} \right) \Big|_0^A = \frac{4C}{\pi A}.$$

Пример 3. В качестве примера НЭ с двузначной СХ рассмотрим реле с гистерезисом (см. рис. 9.6, б). Положим, что $A > b$ и воспользуемся формулами (13.13) и (13.14):

$$\begin{aligned} q(A) &= \frac{1}{\pi A^2} \left[\int_{-A}^{-b} \frac{-2Cx dx}{\sqrt{A^2 - x^2}} + \int_{-b}^b 0 + \int_b^A \frac{2Cx dx}{\sqrt{A^2 - x^2}} \right] = \\ &= \frac{2C}{\pi A^2} \left[\sqrt{A^2 - x^2} \Big|_{-A}^{-b} - \sqrt{A^2 - x^2} \Big|_b^A \right] = \frac{4C}{\pi A^2} \sqrt{A^2 - b^2}; \end{aligned} \quad (13.16)$$

$$q'(A) = \frac{1}{\pi A^2} \left(\int_{-A}^{-b} 0 + \int_{-b}^b -2C dx + \int_b^A 0 \right) = -\frac{4Cb}{\pi A^2}. \quad (13.17)$$

Модуль эквивалентной комплексной характеристики реле с гистерезисом определим по формуле (13.9):

$$R_n(A) = \frac{4C}{\pi A}$$

(он совпадает с соответствующим выражением для идеального реле для $A > b$), а фазовые сдвиги в соответствии с формулой (13.10) равны

$$\varphi_n(A) = -\text{Arctg} \frac{b}{\sqrt{A^2 - b^2}},$$

т.е. реле с гистерезисом вносит отрицательный сдвиг по фазе от $-\pi/2$ при $A = b$ до нуля при $A/b \rightarrow \infty$.

Для сравнения приведем выражения для коэффициентов гармонической линейаризации реле с опережением (рис. 13.4):

$$q(A) = \frac{4C}{\pi A^2} \sqrt{A^2 - b^2}; \quad q'(A) = \frac{4Cb}{\pi A^2}.$$

Отличие от реле с гистерезисом – только в знаке $q'(A)$, т.е. амплитудные характеристики $R_n(A)$ совпадают. Этот элемент вносит положительный фазовый сдвиг по первой гармонике.

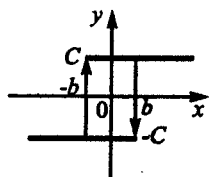


Рис. 13.4. Реле с опережением

13.3.3. Гармоническая линейаризация соединений нелинейных элементов

Если НЭ соединены последовательно (см. рис. 10.2) или два НЭ образуют контур (см. рис. 10.5), то для получения коэффициентов гармонической линейаризации следует вначале построить СХ эквивалентного НЭ $F_3(x)$ и воспользоваться формулами для $q_3(A)$ и $q'_3(A)$, полученными выше.

В том случае, когда НЭ $F_1(x)$ и $F_2(x)$ соединены параллельно, коэффициенты гармонической линейаризации равны сумме коэффициентов гармонической линейаризации, т.е.:

$$q_3(A) = q_1(A) + q_2(A);$$

$$q'_3(A) = q'_1(A) + q'_2(A),$$

что ясно из аддитивности интегралов (13.11).

Для некоторых НЭ со сложными СХ это обстоятельство упрощает получение коэффициентов гармонической линейаризации. Действительно, пусть СХ НЭ имеет вид, изображенный на рис. 13.5, а. На рис. 13.5, б

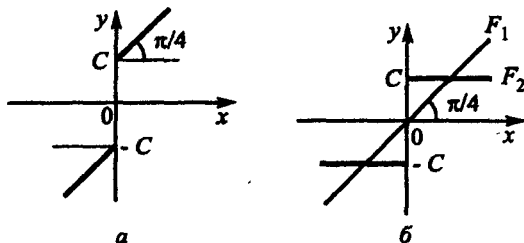


Рис. 13.5. Представление СХ НЭ в виде суммы двух НЭ

показаны две СХ, сумма которых дает исходную СХ. Для этих простейших СХ известны коэффициенты гармонической линеаризации:

$$q_1(A) = 1, \quad q'_1(A) = 0;$$

$$q_2(A) = \frac{4C}{\pi A}, \quad q'_2(A) = 0.$$

Следовательно, для исходного НЭ имеем:

$$q(A) = 1 + \frac{4C}{\pi A}; \quad q'(A) = 0.$$

13.4. ОПРЕДЕЛЕНИЕ ПАРАМЕТРОВ ПЕРИОДИЧЕСКИХ РЕЖИМОВ

Выше было показано, что условия (13.2), (13.5) и (13.6) равносильны тому, что гармонически линеаризованная система (см. рис. 13.1) находится на колебательной границе устойчивости. Поэтому в основу методик определения параметров колебаний A и ω могут быть положены различные критерии устойчивости.

13.4.1. Метод Л. С. Гольдфарба

Графическая методика поиска A^* , ω^* , удовлетворяющих условию (13.6), может быть следующей. Фиксируются значения A_i и строится семейство годографов $W_n(A_i)W(j\omega)$ в функции частоты ω . Если при некотором $A_i = A^*$ годограф пройдет через точку $(-1, j0)$, то в соответствии с критерием Найквиста искомые параметры найдутся немедленно.

По методу Гольдфарба условие (13.6) записывается так:

$$W(j\omega) = -\frac{1}{W_n(A)} = I(A).$$

Строятся два годографа — АФХ ЛЧ $W(j\omega)$ и инверсная характеристика НЭ $I(A)$ (рис. 13.6). Каждый из годографов зависит только от одного искомого параметра. Точка их пересечения отвечает границе устойчивости эквивалентной системы. Можно видеть, что роль критической точки для АФЧ ЛЧ при этом играют точки инверсной характеристики, зависящие от A . Частота колебаний локализуется по меткам на АФХ $W(j\omega)$, а амплитуда — по меткам на инверсной характеристике $I(A)$.

Пример 4. На рис. 13.6 показаны годографы $W(j\omega)$ и $I_1(A)$ для системы (см. рис. 9.14), ЛЧ которой имеет ПФ:

$$W(s) = \frac{1}{s(s+1)^2},$$

а НЭ является идеальным реле:

$$I(A) = -\frac{1}{q(A)} = -\frac{\pi A}{4C}$$

— инверсная характеристика совпадает с отрицательной действительной полуосью. Пересечению кривых отвечают параметры

$$\omega_1^* = 1 \text{ с}^{-1}, \quad A_1^* = 2C/\pi.$$

Изменим условие примера: пусть НЭ является реле с гистерезисом. Ранее были получены коэффициенты гармонической линеаризации (13.16) и (13.17) для этого НЭ. Инверсная характеристика НЭ выражается так:

$$I_2(A) = -\frac{1}{q(A) + jq'(A)} = -\frac{\pi\sqrt{A^2 - b^2}}{4C} - j\frac{\pi b}{4C}.$$

Поскольку мнимая часть постоянна и не зависит от A , график этой характеристики есть прямая в третьем квадранте, параллельная оси абсцисс (см. рис. 13.6). Частота ω_2^* периодического режима здесь меньше, чем ω_1^* , что объясняется дополнительным отрицательным фазовым сдвигом, вносимым НЭ.

Пример 5. Рассмотрим еще один пример, когда НЭ является реле с опережением (см. рис. 13.4). Инверсная характеристика выражается следующим образом:

$$I_3(A) = -\frac{\pi\sqrt{A^2 - b^2}}{4C} + j\frac{\pi b}{4C}.$$

Эта прямая, параллельная оси абсцисс во втором квадранте (см. рис. 13.6). В зависимости от параметров ЛЧ и НЭ здесь могут быть два случая — годографы $W(j\omega)$ и $I_3(A)$ имеют два пересечения или годографы $W(j\omega)$ и $I_3(A)$ не имеют ни одного пересечения (см. рис. 13.6). В последнем случае мы должны говорить об отсутствии периодических режимов гармонической формы. Но здесь могут быть колебания, форма которых сильно отличается от гармонической (например, см. рис. 13.7).

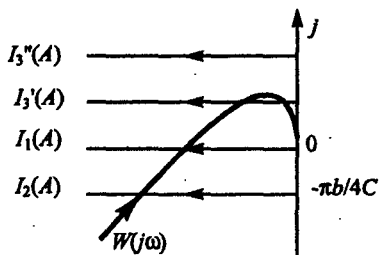


Рис. 13.6. Иллюстрация метода Гольдфарба

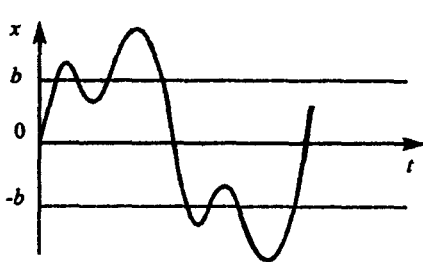


Рис. 13.7. Периодический режим, сильно отличающийся от гармонического

13.4.2. Метод А. А. Вавилова

Условия баланса амплитуд и фаз (13.5) можно записать в логарифмической форме:

$$\begin{aligned} L(\omega) + L_n(A) &= 0; \\ \varphi(\omega) + \varphi_n(A) &= -\pi, \end{aligned} \quad (13.18)$$

где

$$L(\omega) = 20 \lg \text{mod } W(j\omega); \quad L_n(A) = 20 \lg \text{mod } W_n(A).$$

При анализе систем с типовыми НЭ в ряде случаев целесообразно ввести относительную амплитуду гармонического сигнала на входе НЭ

$$\mu = A/b,$$

а комплексный коэффициент передачи представить в виде произведения

$$W_n(\mu) = k_n W_{n0}(\mu),$$

нормирующего коэффициента k_n , зависящего от параметров конкретного НЭ, и нормированного комплексного коэффициента передачи $W_{n0}(\mu)$, единого для всех НЭ данного типа. За постоянную b принимается величина зон нечувствительности, неоднозначности или другой характерный параметр, связанный с осью абсцисс СХ НЭ.

Например, для реле с гистерезисом (см. рис. 9.6, б) и реле с опережением (см. рис. 13.4) нормирующий коэффициент выделяется так:

$$R_n(A) = \frac{4C}{\pi A} \cdot \frac{1}{\mu}; \quad k_n = \frac{4C}{\pi b};$$

при этом фазовые сдвиги в функции μ не зависят от параметров НЭ:

$$\varphi_n(A) = \mp \text{Arctg} \frac{1}{\sqrt{\mu^2 - 1}}.$$

Для идеального реле параметр b условно принимается единичным, т. е. μ численно равняется A и, следовательно, имеем:

$$R_n(\mu) = \frac{4C}{\pi} \cdot \frac{1}{\mu}; \quad k_n = \frac{4C}{\pi}.$$

Нормированные логарифмические характеристики имеют асимптоты, если для относительной амплитуды принята логарифмическая шкала. Например, для идеального реле нормированная логарифмическая характеристика

$$L_{н0}(\mu) = 20 \lg \frac{1}{\mu} = -20 \lg \mu,$$

представляет собой прямую с наклоном -20 дБ/дек. Реле с гистерезисом и опережением для $\mu \geq 1$ имеют ту же характеристику.

Важным результатом нормирования является независимость эквивалентных характеристик данного типа от конкретных значений его параметров. Нормирующий коэффициент k_n при расчете периодических режимов относится к ЛЧ и учитывается сдвигом ЛАЧХ ЛЧ вдоль оси ординат вверх ($k_n > 1$) или вниз ($k_n < 1$). Смещенную на $20 \lg k_n$ ЛАЧХ ЛЧ называют приведенной ЛАЧХ ЛЧ и обозначают $L_0(\omega)$.

Для нормированных характеристик НЭ условия (13.18) запишутся следующим образом:

$$\begin{aligned} L_0(\omega) &= -L_{н0}(\mu); \\ \varphi(\omega) &= -\pi - \varphi_n(\mu) = \varphi_k(\mu). \end{aligned} \quad (13.19)$$

Здесь $L_0(\omega) = 20 \lg k_n + 20 \lg R(\omega)$ — приведенная ЛАЧХ ЛЧ; $L_{н0}(\mu)$ — нормированная амплитудная характеристика НЭ.

Таковы основные положения метода А. А. Вавилова, в котором используются нормированные логарифмические частотные характеристики НЭ и ЛЧ и который базируется на оценке чувствительности периодического решения к высшим гармоникам и малым параметрам. Этот метод обладает существенным достоинством, так как позволяет непосредственно по нормированным логарифмическим характеристикам определять относительную амплитуду и частоту первой гармоники искомого периодического решения [19].

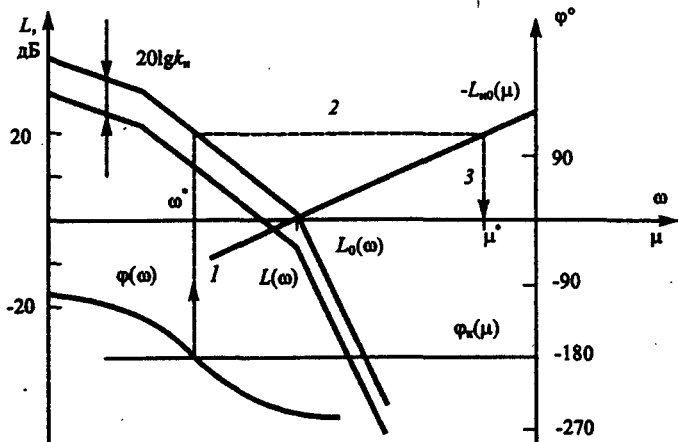


Рис. 13.8. Иллюстрация метода А.А.Вавилова

Пример 6. Рассмотрим иллюстрацию графического метода определения амплитуды $A = \mu b$ и частоты ω первой гармоники периодических колебаний входной переменной НЭ в системе, приведенной на рис. 9.13, для случая безынерционного НЭ с однозначной СХ.

На рис. 13.8 показан пример решения задачи для системы с идеальным релейным элементом и ЛЧ с передаточной функцией

$$W(s) = \frac{k}{s(T_1 s + 1)(T_2 s + 1)}; \quad T_1 > T_2.$$

Нелинейности с однозначными характеристиками не вносят фазового сдвига, поэтому в уравнении баланса фаз (13.19) $\varphi_x(\mu) = -\pi$.

Ход графических построений показан на рис. 13.8 стрелками 1, 2, 3. В рассматриваемой системе существует одно периодическое решение с амплитудой $A^* = \mu^*$ с частотой ω^* первой гармоники.

Примеры графического определения параметров периодического решения в системе с НЭ, имеющим неоднозначную СХ, а также параметров несимметричных периодических режимов приведены в [19].

13.4.3. Метод Е. П. Попова

Рассмотрим систему расчетной структуры с однозначной СХ НЭ, приведенную на рис. 9.13. Как было показано, в этом случае $q'(A) = 0$, т. е. эквивалентная передаточная функция (13.8) совпадает с коэффициентом гармонической линеаризации: $R_n(A) = q(A)$.

Если передаточную функцию ЛЧ записать в виде

$$W(s) = B(s)/D(s),$$

то выражение для характеристического полинома замкнутой гармонически линеаризованной системы получится в виде суммы

$$D_3(s, A) = D(s) + q(A)B(s). \quad (13.20)$$

Гармонически линеаризованная система должна находиться на границе устойчивости. В соответствии с критерием Гурвица предпоследний определитель матрицы Гурвица должен равняться нулю (остальные положительные):

$$\Delta_{n-1}(A) = 0. \quad (13.21)$$

Из условия (13.21) вычисляется значение A^* , при подстановке которого в (13.20) полином D_3 будет иметь пару мнимых корней (остальные левые), т. е.:

$$D_3(j\omega, A^*) = 0. \quad (13.22)$$

Из условия (13.22) следует два уравнения:

$$\operatorname{Re} D_3(j\omega, A^*) = 0;$$

$$\operatorname{Im} D_3(j\omega, A^*) = 0;$$

решение любого из них дает искомую частоту ω^* .

Пример 7. Пусть НЭ представляет собой идеальное реле, для которого ранее получено следующее выражение для коэффициента гармонической линеаризации

$$q(A) = 4C/\pi A.$$

Если ЛЧ системы имеет ПФ

$$W(s) = \frac{1}{s(s+1)^2},$$

то характеристический полином гармонически линеаризованной системы запишется так:

$$D_3(s, A) = s(s+1)^2 + 4C/\pi A = s^3 + 2s^2 + s + 4C/\pi A.$$

Для полинома третьей степени с положительными коэффициентами условие колебательной границы устойчивости сводится к равенству произведений средних и крайних коэффициентов:

$$2 \cdot 1 = 1 \cdot 4C/\pi A.$$

откуда $A^* = 2C/\pi$. Подстановка A^* в $D_3(s, A)$ дает

$$D_3(j\omega, A^*) = -j\omega^3 - 2\omega^2 + j\omega + 2 = 0$$

или

$$\operatorname{Re} D_3(j\omega, A^*) = -2\omega^2 + 2 = 0; \quad \operatorname{Im} D_3(j\omega, A^*) = \omega - \omega^3 = 0.$$

Решая первое уравнение, получаем частоту $\omega^* = 1 \text{ с}^{-1}$ периодического режима. Второе уравнение имеет два корня: $\omega_1 = 1$; $\omega_2 = 0$, последний из которых не принимается во внимание.

13.5. УСТОЙЧИВОСТЬ ПЕРИОДИЧЕСКИХ РЕЖИМОВ

Если периодическое решение уравнения гармонически линеаризованной системы устойчиво, то в нелинейной системе имеют место автоколебания.

Поскольку метод гармонического баланса дает приближенные решения, то и метод исследования устойчивости является приближенным. Идея метода содержится в следующих рассуждениях.

Характеристический полином гармонически линеаризованной системы $D_3(s, A^*)$ имеет пару мнимых корней $\pm j\omega^*$, а остальные — левые (рис. 13.9). При этом колебания переменной на входе НЭ по форме близки к гармонической, их амплитуда близка к A^* , а частота — ω^* .

Если возмущения привели к изменению амплитуды $A^* \pm \Delta A$, то характеристические полиномы $D_3(s, A^* \pm \Delta A)$ уже не имеют чисто мнимых корней. Здесь выделим два случая.

Пусть в первом случае полином $D_3(s, A^* + \Delta A)$ имеет все корни в левой полуплоскости, а полином $D_3(s, A^* - \Delta A)$ имеет пару правых корней (рис. 13.9, а). Тогда периодическое решение устойчиво. Действительно, при увеличении амплитуды гармонически линеаризованная система становится устойчивой — колебания затухают к прежнему зна-

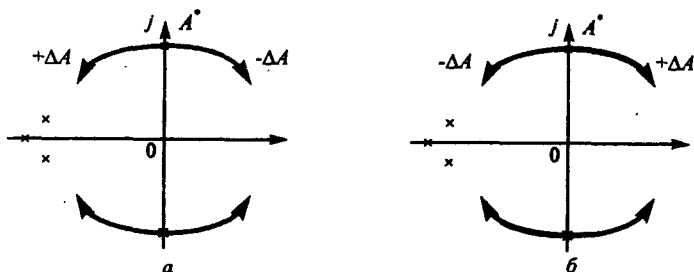


Рис. 13.9. Иллюстрация идеи анализа устойчивости периодического решения

чению A^* . При уменьшении амплитуды из-за возмущений гармонически линеаризованная система становится неустойчивой — колебания расходятся до значения A^* .

Во втором случае полином $D_3(s, A^* + \Delta A)$ имеет пару правых корней, а у полинома $D_3(s, A^* - \Delta A)$ все корни находятся в левой полуплоскости (рис. 13.9, б). Аналогичными рассуждениями приходим к заключению, что в этом случае периодическое решение неустойчиво.

Таким образом, для определения устойчивости периодических решений к полиному $D_3(s, A)$ дважды применяется критерий Гурвица (Рауса) — для $A = A^* + \Delta A$ и $A = A^* - \Delta A$, что составляет завершающую часть метода Е. П. Попова [60].

Метод Л. С. Гольдфарба оперирует годографами на комплексной плоскости — АФХ ЛЧ $W(j\omega)$ и инверсной характеристикой НЭ $I(A)$. Точки инверсной характеристики при конкретных значениях A играют роль критических точек в формулировке критерия Найквиста для устойчивой ЛЧ. Если $A = A^*$, то АФХ ЛЧ проходит через критическую точку: гармонически линеаризованная система находится на колебательной границе устойчивости. Если возмущения привели к изменению амплитуды ($A^* \pm \Delta A$), то могут сложиться две ситуации (рис. 13.10).

В первой ситуации (рис. 13.10, а) увеличение амплитуды приводит к тому, что АФХ ЛЧ не охватывает критическую точку $I(A^* + \Delta A)$: система устойчива, колебания затухают до прежнего значения. При уменьшении амплитуды АФХ ЛЧ охватывает критическую точку $I(A^* + \Delta A)$, а значит, система не устойчива, колебания расходятся. Вывод — периодическое решение устойчиво.

Во второй ситуации (рис. 13.10, б) увеличение амплитуды приводит к охвату АФХ ЛЧ критической точки $I(A^* + \Delta A)$: система становится неустойчивой, колебания расходятся, т. е. амплитуда еще увеличивается. Уменьшение амплитуды, наоборот, приводит к охвату критической точки, система становится устойчивой: колебания затухают;

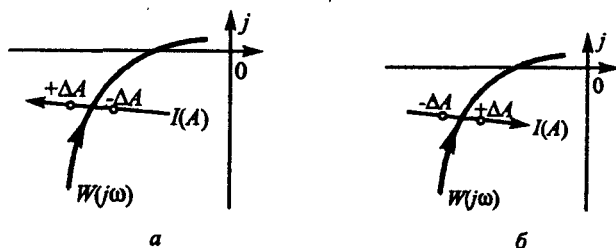


Рис. 13.10. Анализ устойчивости периодического решения методом Гольдфарба

т. е. их амплитуда становится еще меньше. Вывод — периодическое решение неустойчиво.

Обратимся к примерам, показанным на рис. 13.6. Для НЭ типа идеального реле и реле с гистерезисом инверсные годографы $I_1(A)$ и $I_2(A)$ имеют одну точку пересечения с АФХ ЛЧ, причем периодические решения устойчивы. Для НЭ типа реле с опережением инверсная характеристика $I_3(A)$ имеет две точки пересечения с АФХ ЛЧ, причем первое решение (с меньшей частотой) устойчиво, а второе — неустойчиво.

13.6. ЧУВСТВИТЕЛЬНОСТЬ ПЕРИОДИЧЕСКИХ РЕЖИМОВ

При определенных условиях влияние высших гармоник на входе НЭ и малых параметров, обусловленных погрешностью идентификации НЭ и ЛЧ и малыми вариациями параметров системы в процессе эксплуатации, может приводить к существенным изменениям амплитуды и частоты периодического решения. Поэтому при оценке возможности применения метода гармонической линеаризации для исследования периодических режимов в нелинейных системах целесообразно наряду со сглаживающими или резонансными свойствами ЛЧ системы рассматривать чувствительность периодического режима гармонически линеаризованного уравнения системы относительно высших гармоник и малых параметров и условия, при которых высшие гармоники и малые параметры оказывают слабое влияние на точность определения амплитуды и частоты периодического решения.

Как показано А. А. Вавиловым [19], условием применения метода гармонической линеаризации является грубость решения, когда чувствительность периодического решения по отношению к высшим гармоникам и малым параметрам мала, т. е. малым изменениям параметров гармонически линеаризованного уравнения нелинейной системы соответствуют малые отклонения амплитуды и частоты первой гармоники периодического решения.

Решение уравнений баланса амплитуд и фаз (13.5) дает значения A^* и ω^* . Если учесть высшие гармоники и малые параметры, то получают приращение как характеристики линеаризованной системы, так и решения уравнений баланса:

$$R_n(A^* + \delta A)R(\omega^* + \delta\omega) = 1 + \delta R;$$

$$\varphi_n(A^* + \delta A) + \varphi(\omega^* + \delta\omega) = -\pi + \delta\varphi.$$

Точные соотношения между приращениями характеристик и решений $\delta A(\delta R, \delta \varphi)$, $\delta \omega(\delta R, \delta \varphi)$ в общем случае получить нельзя. Поэтому ограничимся рассмотрением приближенных зависимостей.

Пусть в точке исследуемого периодического решения характеристики являются гладкими, а значит, их можно разложить в степенной ряд:

$$R_n(A^*)R(\omega^*) + \frac{\partial R_n(A^*)}{\partial A}R(\omega^*)\delta A + \frac{\partial R(\omega^*)}{\partial \omega}R_n(A^*)\delta \omega + \dots = 1 + \delta R;$$

$$\varphi_n(A^*) + \varphi(\omega^*) + \frac{\partial \varphi_n(A^*)}{\partial A}\delta A + \frac{\partial \varphi(\omega^*)}{\partial \omega}\delta \omega + \dots = -\pi + \delta \varphi.$$

Часть слагаемых в силу (13.5) взаимно уничтожается. Если предположить о малости приращений δA , $\delta \varphi$, то можно пренебречь нелинейными членами разложений. В результате получим приближенную систему линейных уравнений для малых вариаций решений и характеристик:

$$J(A^*, \omega^*) \begin{pmatrix} \delta A \\ \delta \omega \end{pmatrix} = \begin{pmatrix} \delta R \\ \delta \omega \end{pmatrix}, \quad (13.23)$$

где

$$J(A^*, \omega^*) = \begin{pmatrix} \frac{\partial R_n(A^*)}{\partial A}R(\omega^*) & \frac{\partial R(\omega^*)}{\partial \omega}R_n(A^*) \\ \frac{\partial \varphi_n(A^*)}{\partial A} & \frac{\partial \varphi(\omega^*)}{\partial \omega} \end{pmatrix} -$$

— матрица Якоби системы (13.5).

Условием существования решения уравнений (13.23) является невырожденность матрицы

$$\det J(A^*, \omega^*) \neq 0.$$

Решения системы (13.23) имеют вид:

$$\delta A = \left[\frac{\partial \varphi(\omega^*)}{\partial \omega} \delta R - \frac{\partial R(\omega^*)}{\partial \omega} R_n(A^*) \delta \varphi \right] / \det J(A^*, \omega^*) = T_R^A \delta R + T_\varphi^A \delta \varphi;$$

$$\delta \omega = \left[\frac{\partial R_n(A^*)}{\partial A} R(\omega^*) \delta R - \frac{\partial \varphi_n(A^*)}{\partial A} \delta \varphi \right] / \det J(A^*, \omega^*) = T_R^\omega \delta R + T_\varphi^\omega \delta \varphi.$$

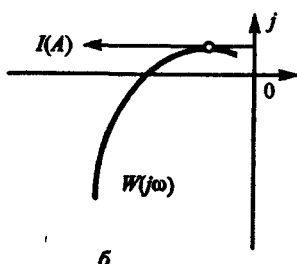
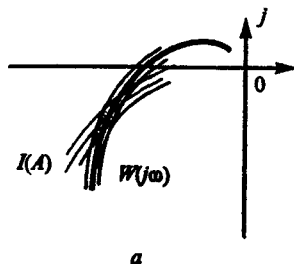


Рис. 13.11. Геометрическая иллюстрация чувствительности периодического решения

Здесь T_R^A , T_φ^A , T_R^∞ , T_φ^∞ — значения функций чувствительности решений к обобщенным вариациям характеристик.

Необходимое условие применимости метода гармонического баланса — функции чувствительности должны быть малы.

Это условие имеет простой геометрический смысл. Применительно к методу Гольдфарба уменьшению чувствительности соответствует более «крутое» пересечение годографов. Если годографы пересекаются под острым углом (см. рис. 13.11, а), то неполная определенность характеристик вызывает большую ошибку локализации решения, даже если ЛЧ имеет хорошие фильтрующие свойства и форма $x(t)$ близка к гармонической. На рис. 13.11, а иллюстрируются вариации характеристик $W(j\omega)$, $I(A)$ в виде полос и вытянутая заштрихованная область, в которой локализованы параметры периодического решения. В том случае, когда годографы только касаются (рис. 13.11, б), якобиан системы (13.5) равен нулю, функции чувствительности стремятся в бесконечность — метод гармонического баланса дает качественную ошибку при исследовании периодического режима.

Если в точке пересечения годографы ортогональны или близки к этому, то метод гармонического баланса может давать хорошие результаты и при отсутствии фильтрующих свойств ЛЧ системы [19].

ЧАСТЬ 5

СИСТЕМЫ УПРАВЛЕНИЯ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

Глава 14

ОБЩИЕ СВЕДЕНИЯ О СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ И ПРОЦЕССАХ В СИСТЕМАХ УПРАВЛЕНИЯ

14.1. СЛУЧАЙНЫЕ ВОЗДЕЙСТВИЯ

Ранее предполагалось, что модель среды на входе системы управления M_F может быть представлена известными детерминированными функциями времени: задающими $g(t)$ и возмущающими $f(t)$ воздействиями. При этом состояние системы в любой момент времени однозначно определяется по известному начальному состоянию. Такие системы с предсказуемым поведением получили название детерминированных.

Однако на практике часто встречаются воздействия *со случайным (стохастическим) характером* изменения, вызывающие в системе появление случайных сигналов. Состояния таких систем могут быть предсказаны только с определенной вероятностью.

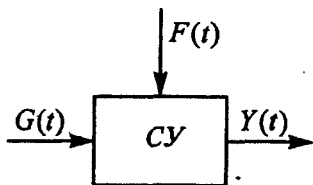


Рис. 14.1. Случайные воздействия, приложенные к системе управления

Случайные воздействия (рис. 14.1), как и детерминированные, разделяют на задающие $G(t)$, несущие полезную информацию для управления (полезный сигнал), и возмущающие $F(t)$ или *помехи*, искажающие полезные сигналы на выходе $Y(t)$ системы управления СУ.

В зависимости от вызывающих их причин помехи могут быть *внешними* и *внутренними*.

Внешние случайные помехи проявляются, например, в виде порывов ветра, действующих на судно, летательный аппарат или радиолокационную антенну; действий морских волн, вызывающих нерегулярную качку судна; изменений напряжения источников питания усилителей и т. д. В экономических системах непредсказуемые политичес-

кие события могут являться внешними помехами, случайным образом изменяющие, например, цены на фондовых биржах. В биологических системах внешними помехами являются случайные изменения окружающей среды, например, лесные пожары или засухи могут оказывать непредвиденное влияние на численность популяций определенных биологических видов. Для приведенных примеров характерно, что амплитудные спектры помех лежат обычно в низкочастотном диапазоне. Внешние помехи, возникающие, например, в радиоприемных устройствах и носящие характер шумов, обусловленных физическими условиями распространения и отражения электромагнитных волн, имеют амплитудные спектры в высокочастотном диапазоне.

Внутренние случайные помехи обусловлены нестабильностью параметров элементов системы. Они могут возникать вследствие флуктуации носителей зарядов в элементах электронных схем, трения в механических соединениях и на границах различных сред и т. д.

Расчет систем управления при случайном характере входных воздействий целесообразно производить с учетом их законов распределения вероятностей или соответствующих им количественных оценок. При этом показатели качества процессов должны отвечать допускаемой вероятности появления случайных воздействий, относительно которых рассчитываются системы. Такой подход заранее предполагает в ряде маловероятных случаев возможность снижения качества работы системы. Для систем с повышенными требованиями, которые не должны снижать качество работы при любых воздействиях, расчет ведется исходя из самых неблагоприятных, но, в то же время, маловероятных для работы системы случаев. При этом используются методы расчета, обычно применяемые для детерминированных систем. Такой подход может приводить к усложнению и удорожанию систем по сравнению с вероятностным подходом.

14.2. СЛУЧАЙНЫЙ ПРОЦЕСС И ЕГО ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

Для расчета систем управления при случайных воздействиях используется математический аппарат теории случайных функций. Случайная функция, как и случайная величина, принимает различные значения в зависимости от исхода опыта ω — элементарного события. Кроме того, случайная функция зависит от некоторого неслучайного параметра λ , т. е. $X(\lambda, \omega)$.

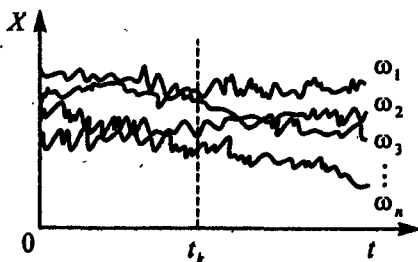


Рис. 14.2. Реализации случайного процесса

Если параметром λ является время t , то случайную функцию называют *случайным процессом* $X(t, \omega)$. При фиксации элементарного события $\omega = \omega_0$ получается регулярная (неслучайная) функция $x(t)$ аргумента t . Конкретный вид случайного процесса при фиксированном ω называется *реализацией* случайного процесса. Совокупность всех возможных реализаций образует *ансамбль* случайного процесса.

На рис. 14.2 приведены реализации случайного процесса $X(t, \omega)$ при различных независимых событиях ω : $\omega_1, \omega_2, \dots, \omega_n, \dots$. Например, реализации могут представлять собой изменения скорости самолета во времени, измеряемой при нескольких полетах, когда приняты все практически возможные меры для поддержания постоянной скорости.

Вероятностный метод исследования случайных процессов ставит задачу не изучения каждой из реализаций $x_i(t)$, входящих в ансамбль, а изучения свойств всего множества в целом с помощью усреднения свойств входящих в него реализаций.

При фиксации параметра $t = t_k$ получается *сечение* случайного процесса, которое зависит только от элементарного события ω и, следовательно, он становится случайной величиной $X(\omega)$.

Если рассматривать n сечений случайного процесса $X(t_1, \omega), X(t_2, \omega), \dots, X(t_n, \omega)$, то в результате получается n -мерная случайная величина, соответствующая случайному процессу $X(t, \omega)$. Обычно для краткости записи аргумент ω опускается.

Случайный процесс $X(t)$ можно рассматривать как случайную величину, изменяющуюся во времени, которая характеризуется многомерной (n -мерной) *функцией распределения вероятностей* (интегральным законом распределения)

$$F_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P\{X(t_1) < x_1; X(t_2) < x_2; \dots; X(t_n) < x_n\}$$

или многомерной *плотностью распределения вероятностей* (дифференциальным законом распределения)

$$W_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

при условии существования частных производных по всем переменным $x_1, x_2, \dots, x_n$. Чем выше порядок n , тем полнее описываются вероятностные свойства случайного процесса.

Часто случайные процессы характеризуют нормальным распределением. Случайный процесс считается *нормально распределенным* или *гауссовым*, если все его плотности вероятности любых порядков нормальны, т. е. распределяются по нормальному закону. Для нормально распределенных случайных процессов их вероятностные свойства полностью определяются двумерной плотностью вероятности:

$$W_2(x_1, x_2; t_1, t_2) = \frac{\partial^2 F_2(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2}.$$

Функция распределения нормального случайного процесса

$$F_2(x_1, x_2; t_1, t_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} W_2(x_1, x_2; t_1, t_2) dx_1 dx_2.$$

Практическое использование многомерных функций распределения затруднительно. Поэтому для упрощения расчета систем управления при случайных воздействиях вводят в рассмотрение более простые, хотя и менее полные количественные характеристики (оценки) случайных процессов, которые, характеризуя случайные процессы, сами по себе являются уже не случайными. Чаще других на практике используют такие характеристики, как математическое ожидание, дисперсия и корреляционная функция.

Математическим ожиданием (средним значением) случайного процесса $X(t)$ называется функция $m_x(t)$, значение которой в каждый момент времени t равно математическому ожиданию соответствующего сечения случайного процесса, т. е.

$$m_x(t) = M[X(t)] = \int_{-\infty}^{\infty} x(t) W_1(x; t) dx.$$

Математическое ожидание представляет собой некоторый усредненный процесс, относительно которого располагаются возможные реализации случайного процесса.

Математическое ожидание регулярного процесса $x(t)$ совпадает с самим процессом, т. е. $m_x(t) = x(t)$.

Дисперсией случайного процесса $X(t)$ называется функция $D_x(t)$, значение которой в каждый момент времени t равно дисперсии соответствующего сечения случайного процесса, т. е.

$$D_x(t) = D[X(t)] = \int_{-\infty}^{\infty} [x(t) - m_x(t)]^2 W_1(x; t) dx.$$

Дисперсия характеризует степень отклонения случайного процесса от его математического ожидания. Дисперсия имеет размерность квадрата случайного процесса, поэтому иногда для удобства расчета используют среднеквадратическое отклонение

$$\sigma_x(t) = \sqrt{D_x(t)},$$

имеющее одинаковую со случайным процессом размерность.

Математическое ожидание и дисперсия задают некоторый «коридор», в котором с определенной вероятностью располагаются возможные реализации случайного процесса, однако эти характеристики не позволяют отразить характер изменения реализаций.

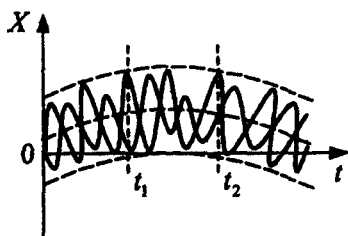
Степень изменчивости или, другими словами, степень вероятностной или статистической зависимости между двумя произвольно выбранными сечениями случайной функции характеризуется корреляционной функцией.

Корреляционной (автокорреляционной) функцией случайного процесса $X(t)$ называется функция двух переменных $K_{xx}(t_1, t_2)$, значение которой для каждой пары произвольно выбранных моментов времени t_1, t_2 равно корреляционному моменту соответствующих сечений *центрированного* случайного процесса $\hat{X}(t) = X(t) - m_x(t)$, т. е.

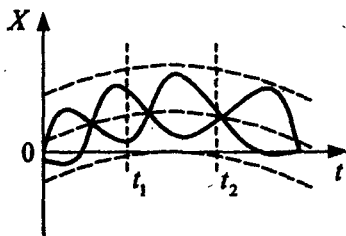
$$\begin{aligned} K_{xx}(t_1, t_2) &= M \left[\hat{X}(t_1) \hat{X}(t_2) \right] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1(t_1) - m_x(t_1)] [x_2(t_2) - m_x(t_2)] W_2(x_1, x_2; t_1, t_2) dx_1 dx_2. \end{aligned}$$

Корреляционная функция $K_{xx}(t_1, t_2)$ отражает вероятностную взаимосвязь (корреляцию) между случайными величинами $X(t_1)$ и $X(t_2)$. Чем меньше вероятностная связь $X(t_1)$ и $X(t_2)$, тем меньше значение $K_{xx}(t_1, t_2)$ и тем быстрее изменяются значения случайного процесса $X(t)$.

На рис. 14.3 изображены реализации двух случайных процессов с одинаковыми математическими ожиданиями и дисперсиями. Случайный процесс на рис. 14.3, а изменяется быстрее, при этом связь между сечениями этого процесса мала. Для случайного процесса на рис. 14.3, б



а



б

Рис. 14.3. Реализации двух случайных процессов с одинаковыми математическими ожиданиями и дисперсиями

вероятностная связь между этими же сечениями больше. Поэтому и корреляционная функция при увеличении разности $t_2 - t_1$ затухает медленнее, чем в случае, приведенном на рис. 14.3, а. Таким образом, можно сказать, что корреляционная функция характеризует внутреннюю структуру случайного процесса.

Корреляционная функция симметрична относительно переменных t_1, t_2 :

$$K_{xx}(t_1, t_2) = K_{xx}(t_2, t_1).$$

Если принять $t_1 = t_2 = t$, то корреляционная функция равна дисперсии, т. е.

$$K_{xx}(t, t) = D_x(t).$$

При исследовании свойств поведения систем управления часто интересуются взаимосвязью нескольких различных случайных процессов системы. Вероятностная связь между двумя случайными процессами $X(t)$ и $Y(t)$ оценивается по их взаимной корреляционной функции.

Взаимной корреляционной функцией двух случайных процессов $X(t)$ и $Y(t)$ называется функция двух переменных $K_{xy}(t_1, t_2)$, значение которой для каждой пары произвольно выбранных моментов времени t_1, t_2 равно корреляционному моменту соответствующих сечений центрированных случайных процессов $\hat{X}(t) = X(t) - m_x(t)$ и $\hat{Y}(t) = Y(t) - m_y(t)$, т. е.

$$\begin{aligned} K_{xy}(t_1, t_2) &= M \left[\hat{X}(t_1) \hat{Y}(t_2) \right] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x(t_1) - m_x(t_1)] [y(t_2) - m_y(t_2)] W_2(x, y; t_1, t_2) dx dy. \end{aligned}$$

Взаимная корреляционная функция несимметрична

$$K_{xy}(t_1, t_2) \neq K_{yx}(t_1, t_2),$$

однако справедливо равенство

$$K_{xy}(t_1, t_2) \neq K_{yx}(t_2, t_1).$$

14.3. СТАЦИОНАРНЫЕ СЛУЧАЙНЫЕ ПРОЦЕССЫ

14.3.1. Определения стационарных случайных процессов

Случайные процессы, протекающие в системах управления, принято разделять на стационарные и нестационарные.

Случайный процесс $X(t)$ называется *стационарным*, если его вероятностные характеристики не зависят от времени, т. е. характеристики случайного процесса $X(t + \tau)$ при любом τ полностью совпадают с соответствующими характеристиками случайного процесса $X(t)$.

Различают стационарность процессов в узком и в широком смысле.

Случайный процесс $X(t)$ называется *стационарным в узком смысле* или *строго стационарным*, если его n -мерные функции распределения и плотности вероятности при любом n не зависят от временного сдвига τ для всех моментов $t_1, t_2, \dots, t_n$, т. е.

$$F_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = F_n(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau),$$

$$W_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = W_n(x_1, x_2, \dots, x_n; t_1 + \tau, t_2 + \tau, \dots, t_n + \tau), \quad n = 1, 2, \dots$$

Из приведенного определения следует инвариантность от времени всех вероятностных характеристик стационарного случайного процесса. Вероятностные характеристики стационарного процесса не зависят от начала отсчета во времени.

Случайный процесс $X(t)$ называется *стационарным в широком смысле* или *слабо стационарным*, если его математическое ожидание постоянно

$$M[X(t)] = m_x(t) = m_x = \text{const},$$

корреляционная функция зависит только от одной переменной $\tau = t_2 - t_1$

$$\begin{aligned} K_{xx}(\tau) &= K_{xx}(t_1, t_1 + \tau) = M\left[\dot{X}(t_1)\dot{X}(t_1 + \tau)\right] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1(t_1) - m_x][x_2(t_1 + \tau) - m_x] W_2(x_1, x_2; \tau) dx_1 dx_2. \end{aligned}$$

Случайные процессы $X(t)$ и $Y(t)$ называются *стационарно связанными*, если их взаимная корреляционная функция зависит только от временного сдвига $\tau = t_2 - t_1$, т. е.

$$K_{xy}(t_1, t_2) = K_{xy}(\tau).$$

Случайные процессы, стационарные в узком смысле, обязательно стационарны и в широком смысле; однако обратное утверждение неверно. Для нормальных случайных процессов понятия стационарности в широком и узком смысле совпадают, поскольку для них математическое ожидание и корреляционная функция полностью определяют n -мерную плотность вероятности. В дальнейшем будут рассматриваться только случайные функции, стационарные в широком смысле.

Если случайный процесс не удовлетворяет условиям стационарности, то он называется *нестационарным*. Переходный процесс не является стационарным.

При изучении стационарных случайных процессов в системах управления используют два понятия средних значений: по множеству и по времени.

Среднее значение по множеству (среднее по ансамблю) определяется по множеству реализаций случайного процесса в один и тот же момент времени

$$m_x = \int_{-\infty}^{\infty} x W_1(x) dx.$$

Величина m_x представляет собой операцию вероятностного усреднения случайной величины X (случайного процесса $X(t)$ в определенный момент времени), при котором каждое возможное значение x случайной величины принимается с весом, равным элементу вероятности $W_1(x)dx$.

Среднее значение по времени определяется по одной реализации $x(t)$ случайного процесса на всем бесконечном интервале времени T , т. е.

$$\bar{x} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t) dt$$

при условии существования предела.

В общем случае среднее значение по времени различно для отдельных реализаций случайного процесса $X(t)$. Кроме того, для одного и того же случайного процесса средние значения по множеству и по времени различны.

Например, стационарный случайный процесс, представленный на рис. 14.4 отдельными реализациями, характеризуется тем, что среднее

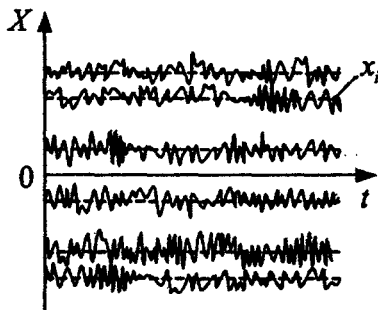


Рис. 14.4. Стационарный случайный процесс

по множеству $m_x = 0$ в любой момент времени, тогда как для какой-либо реализации x_i среднее по времени $\bar{x}_i \neq 0$.

14.3.2. Эргодические случайные процессы

Существуют стационарные случайные процессы, для которых средние значения по множеству и по времени равны, т. е. $m_x = \bar{x}$. Такие случайные процессы называются эргодическими.

Случайный процесс $X(t)$ называется *эргодическим*, если его вероятностные характеристики не зависят от номера реализации.

Один и тот же случайный процесс может быть эргодическим по отношению к одним вероятностным характеристикам и неэргодическим по отношению к другим. Стационарный случайный процесс $X(t)$ эргодичен, если его корреляционная функция отвечает достаточно-му условию

$$\lim_{|\tau| \rightarrow \infty} K_{xx}(\tau) = 0.$$

Свойство эргодичности имеет важное практическое значение. Для большинства систем экспериментально получить множество всех возможных реализаций случайного процесса в один и тот же момент времени для определения среднего значительно сложнее, чем осуществить наблюдение за отдельной реализацией на достаточно длительном интервале времени.

Дисперсия эргодического случайного процесса

$$D_x = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - \bar{x}]^2 dt,$$

где $x(t)$ — любая реализация случайного процесса.

$$K_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - \bar{x}][x(t + \tau) - \bar{x}] dt;$$

если $\bar{x} = 0$, то

$$K_{xx}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t)x(t + \tau) dt.$$

Начальное значение корреляционной функции $K_{xx}(\tau)$ равно дисперсии случайного процесса: $K_{xx}(0) = D_x$.

Взаимная корреляционная функция эргодических случайных процессов

$$K_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T [x(t) - \bar{x}][y(t + \tau) - \bar{y}] dt,$$

где $x(t), y(t)$ — любые реализации случайных процессов $X(t)$ и $Y(t)$; $\bar{x}, \bar{y}$ — средние значения по времени. Из этого выражения следует, что

$$K_{xy}(\tau) = K_{yx}(-\tau), \quad K_{yx}(\tau) = K_{xy}(-\tau).$$

Функция $K_{xy}(\tau)$ характеризует вероятностную взаимосвязь двух случайных процессов в системе $X(t)$ и $Y(t)$ в различные моменты времени, отстоящие друг от друга на интервал времени τ . При $\tau = 0$ значение $K_{xy}(0)$ устанавливает эту взаимосвязь в один и тот же момент времени.

Для случайных процессов $X(t)$ и $Y(t)$, вероятностно не связанных между собой, взаимная корреляционная функция

$$K_{xy}(\tau) \equiv 0.$$

14.3.3. Основные свойства корреляционной функции

К числу основных свойств корреляционной функции можно отнести следующие.

1. Начальное значение

$$K_{xx}(0) = D_x.$$

2. Ограниченность

$$|K_{xx}(\tau)| \leq D_x.$$

3. Четность

$$K_{xx}(\tau) = K_{xx}(-\tau).$$

Отсюда следует, что график корреляционной функции симметричен относительно оси ординат.

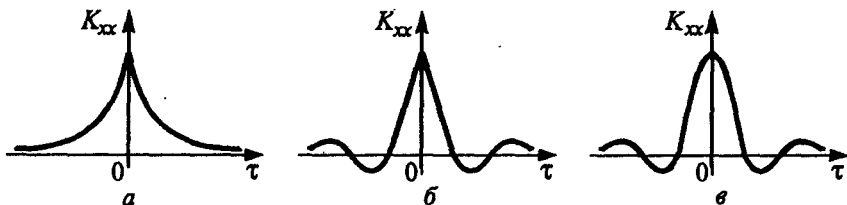


Рис. 14.5. Типичные корреляционные функции

4. Для суммы случайных процессов $Z(t) = X(t) + Y(t)$:

$$K_{zz}(\tau) = K_{xx}(\tau) + K_{xy}(\tau) + K_{yx}(\tau) + K_{yy}(\tau).$$

5. Для постоянного процесса $x(t) = a$:

$$K_{xx}(\tau) = a^2,$$

т. е. корреляционная функция также постоянна.

6. Для гармонического процесса $x(t) = A \sin(\omega_0 t + \varphi)$:

$$K_{xx}(\tau) = 0,5A^2 \cos \omega_0 \tau,$$

т. е. корреляционная функция представляет собой четную относительно τ гармоническую функцию той же частоты ω_0 , но не зависящую от сдвига фазы φ .

Примеры типичных корреляционных функций стационарных случайных процессов представлены на рис. 14.5. Часто их аппроксимируют аналитическими выражениями:

$$K_{xx}(\tau) = D e^{-\mu|\tau|} \quad (\text{рис. 14.5, а}); \quad K_{xx}(\tau) = D e^{-\mu|\tau|} \cos \omega_0 \tau \quad (\text{рис. 14.5, б});$$

$$K_{xx}(\tau) = D e^{-\mu|\tau|} \cos(\omega_0 \tau + \gamma \sin \omega_0 |\tau|), \quad (\text{рис. 14.5, в}),$$

где D — дисперсия случайного процесса; μ , ω_0 , γ — константы.

Из рис. 14.5 видно, что с увеличением τ взаимосвязь между $X(t)$ и $X(t + \tau)$ ослабевает и корреляционная функция уменьшается. Наличие отрицательных значений корреляционных функций (см. рис. 14.5, б, в) указывает на то, что в структуре случайного процесса присутствует периодичность, в связи с чем наблюдается отрицательная корреляция между значениями случайного процесса: положительным отклонениям от среднего в одном сечении соответствуют через определенный промежуток времени отрицательные отклонения в другом сечении, и наоборот. Чем выше изменчивость случайного процесса, т. е. чем большими значениями производной он характеризуется, или, другими словами, чем более высокочастотные составляющие присутствуют в спектре про-

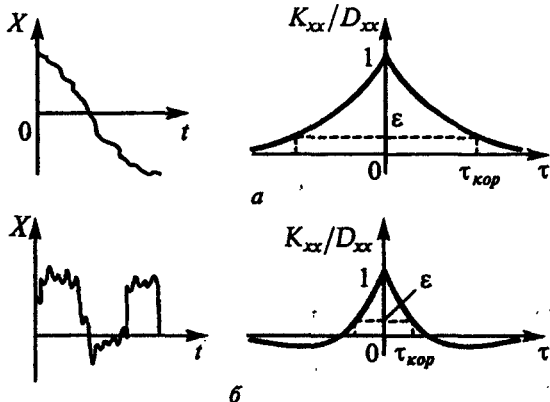


Рис. 14.6. Случайные процессы и их корреляционные функции

процесса, тем быстрее убывает соответствующая ему корреляционная функция (рис. 14.6, а, б).

Мерой длительности существования корреляционной связи между двумя сечениями $X(t)$ и $X(t + \tau)$ служит величина $\tau_{\text{кор}}$, называемая *временем (интервалом) корреляции* случайного процесса. При этом выполняется условие

$$|K_{xx}(\tau_{\text{кор}}) / D_{xx}| \leq \varepsilon,$$

где ε — малая величина (обычно на практике принимают $\varepsilon = 0,05 - 0,1$).

Случайный процесс, в котором отсутствует взаимосвязь между предыдущими и последующими значениями, называется *белым шумом* или *абсолютно случайным процессом*. Это свойство выделяет его среди других случайных процессов, определяет его особую роль как простейшей модели случайного процесса. Для белого шума $\tau_{\text{кор}} = 0$. Вид белого шума и его корреляционной функции приведены на рис. 14.7.

Корреляционная функция белого шума описывается δ -функцией

$$K_{xx}(\tau) = \bar{N} \delta(\tau),$$

где $\bar{N}$ — интенсивность.

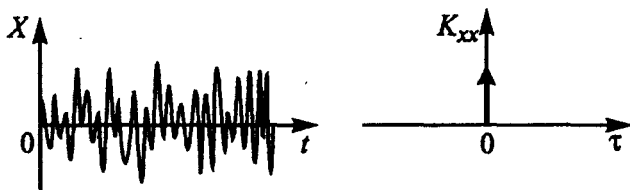


Рис. 14.7. Белый шум и его корреляционная функция

Случайный процесс типа белого шума физически не осуществим, поскольку его дисперсия $D_x = K_{xx}(0)$ принимает бесконечно большое значение. Интенсивность белого шума принимается в качестве его характеристики вместо дисперсии, которая всегда бесконечна.

14.3.4. Спектральная плотность и ее связь с корреляционной функцией

Преобразование Фурье корреляционной функции $K_{xx}(\tau)$ случайного процесса $X(t)$ называется его *спектральной плотностью* $S_{xx}(\omega)$, т. е.

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} K_{xx}(\tau) e^{-j\omega\tau} d\tau. \quad (14.1)$$

Учитывая формулу Эйлера (L. Euler) $e^{-j\omega\tau} = \cos\omega\tau - j\sin\omega\tau$, соотношение (14.1) можно записать в виде

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} K_{xx}(\tau) \cos\omega\tau d\tau - j \int_{-\infty}^{\infty} K_{xx}(\tau) \sin\omega\tau d\tau.$$

Подынтегральные выражения представляют собой: $K_{xx}(\tau)\cos\omega\tau$ — четную функцию; $K_{xx}(\tau)\sin\omega\tau$ — нечетную функцию, поэтому

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} K_{xx}(\tau) \cos\omega\tau d\tau = 2 \int_0^{\infty} K_{xx}(\tau) \cos\omega\tau d\tau. \quad (14.2)$$

Спектральная плотность является действительной и четной функцией относительно частоты ω , т. е.

$$S_{xx}(\omega) = S_{xx}(-\omega).$$

Следовательно, график $S_{xx}(\omega)$ симметричен относительно оси ординат.

Для любого стационарного случайного процесса спектральная плотность представляет собой неотрицательную функцию ω .

По известной спектральной плотности $S_{xx}(\omega)$ случайного процесса можно определить его корреляционную функцию, используя обратное преобразование Фурье:

$$K_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) e^{j\omega\tau} d\omega = \frac{1}{\pi} \int_0^{\infty} S_{xx}(\omega) \cos\omega\tau d\omega, \quad (14.3)$$

откуда получим выражение для дисперсии случайного процесса

$$D_x = K_{xx}(0) = \frac{1}{\pi} \int_0^{\infty} S_{xx}(\omega) d\omega \quad (14.4)$$

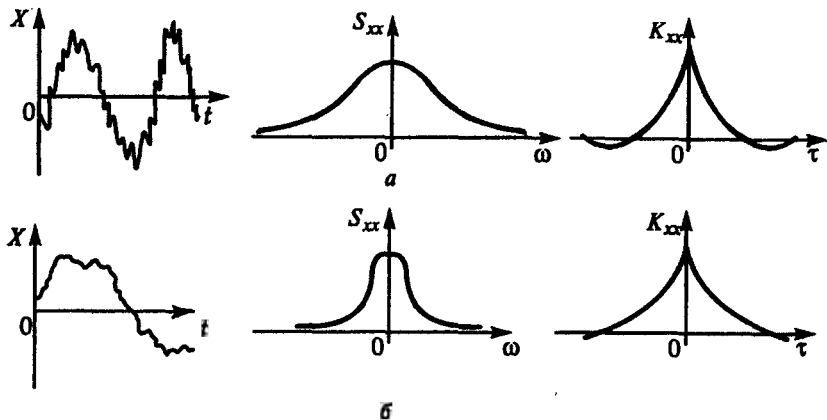


Рис. 14.8. Случайные процессы, их спектральные плотности и корреляционные функции

Соотношения (14.2) и (14.3), связывающие корреляционную функцию и спектральную плотность стационарного случайного процесса, называются *формулами Винера (N. Wiener)-Хинчина*.

Графики спектральной плотности и корреляционной функции в зависимости от характера случайного процесса приведены на рис. 14.8, из которого видно, что чем шире график $S_{xx}(\omega)$, тем уже график $K_{xx}(\tau)$. Это объясняется тем, что случайный процесс, соответствующий графикам на рис. 14.8, а, содержит более высокочастотные, а следовательно, более изменчивые составляющие по сравнению со случайным процессом, представленным на рис. 14.8, б.

Важное практическое значение понятия спектральной плотности во многом определяется тем, что с его помощью сравнительно просто изучается прохождение случайных сигналов через линейные системы.

Поясним физический смысл спектральной плотности. Для этого введем в рассмотрение *финитную функцию*, которая имеет описание вида:

$$x_T(t) = \begin{cases} x(t) & \text{при } |t| \leq T; \\ 0 & \text{при } |t| > T, \end{cases}$$

где T — интервал длительности функции $x(t)$.

Применяя преобразование Фурье к этой функции, получаем спектральную характеристику

$$X_T(j\omega) = \int_{-\infty}^{\infty} x_T(t) e^{-j\omega t} dt = \int_{-T}^T x(t) e^{-j\omega t} dt.$$

Тогда, в соответствии с формулой Парсеваля (см. п. 3.11.2) для финитной функции справедливы равенства:

$$\int_{-\infty}^{\infty} x_T^2(t) dt = \int_{-T}^T x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X_T(j\omega)|^2 d\omega,$$

причем подынтегральное выражение $|X_T(j\omega)|^2$ представляет собой *спектральную плотность энергии функции $x_T(t)$* . После деления на $2T$ и устремления интервала длительности T к бесконечности получаем

$$\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x^2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(j\omega)|^2 d\omega.$$

При $x = 0$ левая часть равенства равна дисперсии, тогда

$$D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(j\omega)|^2 d\omega = \frac{1}{\pi} \int_{-\infty}^{\infty} \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(j\omega)|^2 d\omega.$$

Сравнивая это выражение с ранее полученным (14.4), можно записать

$$S_{xx}(\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(j\omega)|^2 = \lim_{T \rightarrow \infty} \frac{1}{2T} X_T(j\omega) X_T(-j\omega).$$

Таким образом, спектральная плотность представляет собой предел отношения спектральной плотности энергии случайного процесса, выражаемого финитной функцией, к интервалу длительности при бесконечном его увеличении. Следовательно, *спектральная плотность определяет спектральную мощность случайного процесса*.

Преобразование Фурье взаимной корреляционной функции $K_{xy}(\tau)$ двух случайных процессов $X(t)$ и $Y(t)$ называется их *взаимной спектральной плотностью $S_{xy}(j\omega)$* , т. е.

$$S_{xy}(j\omega) = \int_{-\infty}^{\infty} K_{xy}(\tau) e^{-j\omega\tau} d\tau.$$

Взаимная спектральная плотность $S_{xy}(j\omega)$ является комплексной функцией и представляет собой меру вероятностной связи двух случайных процессов $X(t)$ и $Y(t)$.

Для взаимной спектральной плотности получаем

$$S_{xy}(j\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} X_T(j\omega) Y_T(-j\omega).$$

Знание спектральной плотности позволяет определить взаимную корреляционную функцию случайных процессов $X(t)$ и $Y(t)$:

$$K_{xy}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xy}(j\omega) e^{j\omega\tau} d\omega.$$

Пример. Корреляционная функция стационарного случайного процесса $X(t)$ задается формулой

$$K_{xx}(\tau) = D e^{-\mu|\tau|}.$$

Требуется определить спектральную плотность $S_{xx}(\omega)$.

В соответствии с (14.1) можно записать:

$$\begin{aligned} S_{xx}(\omega) &= D \int_{-\infty}^{\infty} e^{-\mu|\tau|} e^{-j\omega\tau} d\tau = D \left[\int_{-\infty}^0 e^{\mu\tau} e^{-j\omega\tau} d\tau + \int_0^{\infty} e^{-\mu\tau} e^{-j\omega\tau} d\tau \right] = \\ &= D \left[\frac{-1}{\mu - j\omega} e^{-(\mu - j\omega)\tau} + \frac{-1}{\mu + j\omega} e^{-(\mu + j\omega)\tau} \right]_0^{\infty} = D \left[\frac{1}{\mu - j\omega} + \frac{1}{\mu + j\omega} \right] = \frac{2D\mu}{\mu^2 + \omega^2}. \end{aligned}$$

При уменьшении параметра μ в спектре случайного процесса большие удельные вес приобретают низкочастотные составляющие: кривая спектральной плотности вытягивается вверх, одновременно сжимаясь с боков. В пределе при $\mu \rightarrow 0$ кривая $S_{xx}(\omega)$ вырождается в вертикальную линию. При увеличении параметра μ преобладание низкочастотных составляющих в спектре случайного процесса уменьшается; кривая $S_{xx}(\omega)$ становится более пологой вдоль оси частот.

14.4. МОДЕЛИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЙ

Рассмотрим некоторые широко распространенные модели случайных воздействий, представленных своими вероятностными характеристиками.

1. *Идеальный белый шум.* Случайное воздействие, характеризующееся постоянной спектральной плотностью во всем диапазоне частот $(-\infty, \infty)$:

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} \bar{N} \delta(\tau) e^{-j\omega\tau} d\tau = \bar{N} e^{-j\omega\tau} \Big|_{\tau=0} = \bar{N} = c^2,$$

где c^2 — интенсивность. График спектральной плотности белого шума представлен на рис. 14.9. Следовательно, спектральная мощность идеального белого шума равномерно распределена по всем частотам.

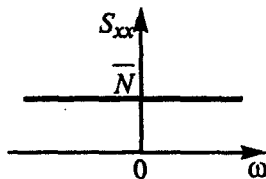


Рис. 14.9. Спектральная плотность идеального белого шума

Корреляционная функция

$$K_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) e^{j\omega\tau} d\omega = \frac{1}{\pi} \int_0^{\infty} S_{xx}(\omega) \cos \omega\tau d\omega = c^2 \delta(\tau).$$

График $K_{xx}(\tau)$ приводился на рис. 14.7.

2. *Реальный белый шум.* Поскольку модель в виде белого шума является идеализацией воздействия, на практике нередко используется

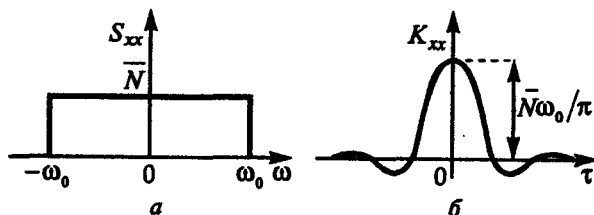


Рис. 14.10. Спектральная плотность (а) и корреляционная функция (б) реального белого шума

более реалистичная модель воздействия в виде реального белого шума с ограниченной спектральной плотностью (рис. 14.10, а):

$$S_{xx}(\omega) = \bar{N} = c^2 \text{ при } |\omega| \leq \omega_0; S_{xx}(\omega) = 0 \text{ при } |\omega| > \omega_0.$$

Дисперсия в этом случае будет равна

$$D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) d\omega = \frac{c^2 \omega_0}{\pi},$$

а корреляционная функция (рис. 14.10, б).

$$K_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} c^2 e^{j\omega\tau} d\omega = \frac{c^2}{2\pi} \frac{1}{j\tau} e^{j\omega\tau} \Big|_{-\omega_0}^{\omega_0} = \frac{c^2}{\pi\tau} \sin \omega_0 \tau.$$

3. *Гладкий реальный белый шум.* При аналитических исследованиях более предпочтительной является аппроксимация спектральной плотности идеального белого шума гладкой функцией без разрывов:

$$S_{xx}(\omega) = \frac{2D\mu}{\omega^2 + \mu^2}, \quad (14.5)$$

где D, μ — параметры характеристики, причем параметр затухания μ имеет такую же единицу измерения как и частота ω , т. е. рад/с. График спектральной плотности $S_{xx}(\omega)$ приведен на рис. 14.11, а.

В интервале $\omega \in [-\mu, \mu]$ эта функция с приемлемой для практики точностью аппроксимирует идеальный белый шум.

Выражение для дисперсии в этом случае будет иметь вид

$$D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2D\mu}{\omega^2 + \mu^2} d\omega = \frac{D\mu}{\pi} \frac{1}{\mu} \arctg \frac{\omega}{\mu} \Big|_{-\infty}^{\infty} = D,$$

а для корреляционной функции (рис. 14.11, б)

$$K_{xx}(\tau) = D e^{-\mu|\tau|}.$$

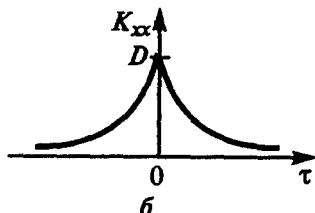
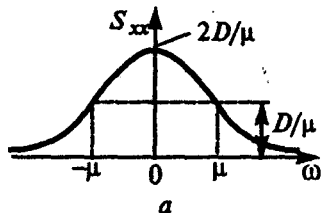


Рис. 14.11. Спектральная плотность (а) и корреляционная функция (б) гладкого реального белого шума

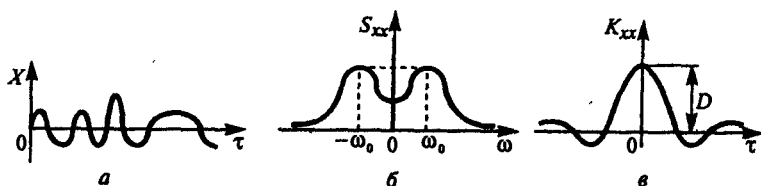


Рис. 14.12. Случайный процесс типа нерегулярной качки и его вероятностные характеристики

4. Нерегулярная качка. Многие движущиеся объекты, например, летательные аппараты, суда, автомобили находятся под влиянием нерегулярных воздействий (атмосферные возмущения, волнение моря, неровная дорога). Их движение осуществляется по случайному закону (рис. 14.12, а). Такое случайное движение объектов называется нерегулярной качкой в отличие от регулярной качки, представляющей собой периодическое движение.

Корреляционная функция нерегулярной качки

$$K_{xx}(\tau) = D e^{-\mu|\tau|} \cos \omega_0 \tau,$$

спектральная плотность

$$S_{xx}(\omega) = \mu D \left[\frac{1}{\mu^2 + (\omega_0 - \omega)^2} + \frac{1}{\mu^2 + (\omega_0 + \omega)^2} \right],$$

где ω_0 — резонансная частота, μ — параметр затухания, D — дисперсия.

Рассмотрим далее спектральные плотности некоторых регулярных процессов, используемых в качестве моделей воздействий.

5. Постоянный процесс $x(t) = a$. Спектральная плотность

$$S_{xx}(\omega) = 2\pi a^2 \delta(\omega).$$

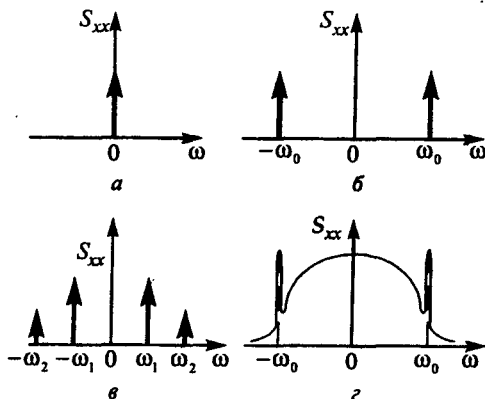


Рис. 14.13. Спектральные плотности различных случайных процессов

На рис. 14.13, а представлен график спектральной плотности, из которого следует, что вся мощность процесса сосредоточена на нулевой частоте.

6. Гармонический процесс $x(t) = A \sin(\omega_0 t + \varphi)$. Спектральная мощность

$$S_{xx}(\omega) = 2\pi \left(\frac{A}{2} \right)^2 [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)].$$

Из графика $S_{xx}(\omega)$, изображенного на рис. 14.13, б, видно, что вся мощность процесса сосредоточена на частотах $-\omega_0$ и ω_0 . Если рассматривать только положительный диапазон частот, то мощность процесса будет сосредоточена на частоте ω_0 .

7. Частичная сумма ряда Фурье $x(t) = c_0 + \sum_{k=1}^n c_k \sin(\omega_k t + \varphi_k)$. Выражение для спектральной мощности имеет вид

$$S_{xx}(\omega) = 2\pi \left\{ c_0^2 \delta(\omega) + \sum_{k=1}^n \left(\frac{c_k}{2} \right)^2 [\delta(\omega - \omega_k) + \delta(\omega + \omega_k)] \right\}.$$

На рис. 14.13, в приведен график $S_{xx}(\omega)$, имеющий вид линейчатого (дискретного) спектра. При этом импульсы условно изображаются таким образом, что их амплитуды пропорциональны квадратам соответствующих коэффициентов Фурье, т. е. c_0^2 и $(c_k/2)^2$.

8. Случайные процессы, на которые накладываются регулярные гармонические составляющие, имеют спектральную плотность с выраженными резонансными пиками, соответствующими скрытым гармоникам, трудно обнаруживаемым в отдельных реализациях (рис. 14.13, г).

АНАЛИЗ ЛИНЕЙНЫХ СИСТЕМ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

Процессы в системе управления определяются начальными условиями и воздействиями. В общем случае начальные условия можно считать случайными величинами, а воздействия — случайными функциями времени. Поэтому и процессы в системе являются случайными функциями времени.

Достаточно полными характеристиками случайных процессов в системе являются их многомерные функции распределения. Для многих систем управления случайные начальные условия и случайные воздействия подчиняются нормальному закону распределения.

Многомерные нормальные законы распределения случайных процессов в системе полностью определяются математическими ожиданиями, автокорреляционными и взаимно корреляционными функциями этих процессов. Если же процессы в системе нельзя приближенно считать нормально распределенными, то для более полного представления о случайных процессах привлекаются дополнительные характеристики, например, эксцесс и асимметрия, характеризующие отличие закона распределения случайного процесса от нормального.

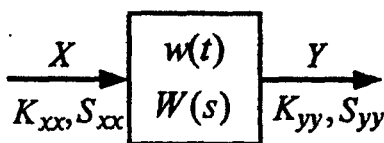
При исследовании устойчивых систем управления, находящихся в установившихся условиях приложения случайных воздействий, рассматривают как установившиеся, так и переходные процессы. Установившиеся процессы, в отличие от переходных, не зависят от начальных условий и определяются только воздействиями. Это означает, что воздействия были приложены к системе в момент времени $t = -\infty$, т. е. действуют бесконечно долго.

Задача анализа состоит в определении вероятностных характеристик реакции звена, объекта и замкнутой системы при подаче случайных воздействий на их входы. Будем исследовать установившиеся режимы, в предположении, что переходные процессы полностью завершены. Все рассматриваемые стационарные случайные процессы отвечают условию эргодичности.

15.1. ПРЕОБРАЗОВАНИЕ СЛУЧАЙНОГО СИГНАЛА ЛИНЕЙНЫМ ЗВЕНОМ

Рассмотрим преобразование случайного сигнала $X(t)$ при его прохождении через линейное звено (рис. 15.1) с импульсной переходной функцией (функцией веса)

$$w(t) = L^{-1}\{W(s)\} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} W(s)e^{st} ds,$$



где c — абсцисса сходимости, $W(s)$ — передаточная функция.

Предполагается, что случайный сигнал $X(t)$ на входе звена стационарный, характеризуется нулевым средним значением $m_x = 0$, корреляционной функцией $K_{xx}(\tau)$ и спектральной плотностью $S_{xx}(\omega)$. Если полюсы передаточной функции $W(s)$ левые, то установившийся случайный сигнал $Y(t)$ на выходе звена также стационарный со средним значением $m_y = 0$. Однако характеристики $K_{yy}(\tau)$ и $S_{yy}(\omega)$ при этом изменяются.

Для установления связи между вероятностными характеристиками входного $X(t)$ и выходного $Y(t)$ сигналов можно воспользоваться формулой свертки

$$y(t) = \int_{-\infty}^{\infty} w(t - \xi)x(\xi)d\xi = \int_{-\infty}^{\infty} w(\xi)x(t - \xi)d\xi,$$

где ξ — независимая переменная интегрирования.

Для момента времени $t + \tau$ имеем:

$$y(t + \tau) = \int_{-\infty}^{\infty} w(\eta)x(t + \tau - \eta)d\eta,$$

где η — другое обозначение независимой переменной интегрирования. Тогда корреляционная функция выходного сигнала

$$\begin{aligned} K_{yy}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(t)y(t + \tau)dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[\int_{-\infty}^{\infty} w(\xi)x(t - \xi)d\xi \int_{-\infty}^{\infty} w(\eta)x(t + \tau - \eta)d\eta \right] dt. \end{aligned}$$

Меняя порядок интегрирования, получаем

$$K_{yy}(\tau) = \int_{-\infty}^{\infty} w(\xi)d\xi \int_{-\infty}^{\infty} w(\eta) \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t - \xi)x(t + \tau - \eta)dt \right] d\eta.$$

Если ввести обозначение

$$K_{xx}(\xi + \tau - \eta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t - \xi)x(t + \tau - \eta)dt,$$

то можно записать интегральное соотношение

$$K_{yy}(\tau) = \int_{-\infty}^{\infty} w(\xi) d\xi \int_{-\infty}^{\infty} w(\eta) K_{xx}(\xi + \tau - \eta) d\eta. \quad (15.1)$$

связывающее корреляционную функцию $K_{xx}(\tau)$ сигнала на входе звена, его импульсную переходную характеристику $w(t)$ с корреляционной функцией сигнала $K_{yy}(\tau)$ на выходе.

Дисперсия случайного сигнала на выходе

$$D_y = K_{yy}(0) = \int_{-\infty}^{\infty} w(\xi) d\xi \int_{-\infty}^{\infty} w(\eta) K_{xx}(\xi - \eta) d\eta.$$

В частности, когда на входе действует стационарный случайный сигнал с корреляционной функцией $K_{xx}(\tau) = c^2 \delta(\tau)$, соответствующей идеальному белому шуму, дисперсия сигнала на выходе

$$D_y = \int_{-\infty}^{\infty} w(\xi) d\xi \int_{-\infty}^{\infty} w(\eta) [c^2 \delta(\xi - \eta)] d\eta = c^2 \int_{-\infty}^{\infty} w^2(\xi) d\xi.$$

Взаимная корреляционная функция сигналов $Y(t)$ и $X(t)$:

$$\begin{aligned} K_{yx}(\tau) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T y(t) x(t - \tau) dt = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[\int_{-\infty}^{\infty} w(\xi) x(t - \xi) d\xi \right] x(t - \tau) dt = \\ &= \int_{-\infty}^{\infty} w(\xi) d\xi \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T x(t - \xi) x(t - \tau) dt = \int_{-\infty}^{\infty} w(\xi) K_{xx}(\tau - \xi) d\xi. \end{aligned}$$

Перейдем к установлению связи между спектральными плотностями сигналов $Y(t)$ и $X(t)$. Спектральная плотность сигнала на выходе звена имеет вид

$$S_{yy}(\omega) = \int_{-\infty}^{\infty} K_{yy}(\tau) e^{-j\omega\tau} d\tau. \quad (15.2)$$

Подставив выражение (15.1) в (15.2), получаем

$$\begin{aligned} S_{yy}(\omega) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} w(\xi) d\xi \int_{-\infty}^{\infty} w(\eta) K_{xx}(\xi + \tau - \eta) d\eta \right] e^{-j\omega\tau} d\tau = \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} w(\xi) d\xi \int_{-\infty}^{\infty} w(\eta) K_{xx}(\xi + \tau - \eta) e^{-j\omega(\xi + \tau - \eta)} e^{j\omega(\xi - \eta)} d\eta \right] d\tau = \\ &= \int_{-\infty}^{\infty} w(\eta) e^{-j\omega\eta} d\eta \int_{-\infty}^{\infty} w(\xi) e^{j\omega\xi} d\xi \int_{-\infty}^{\infty} K_{xx}(\xi + \tau - \eta) e^{-j\omega(\xi + \tau - \eta)} d\tau. \end{aligned}$$

Преобразование Фурье весовой функции $w(t)$ представляет собой комплексный коэффициент передачи $W(j\omega)$, причем

$$W(j\omega) = \int_{-\infty}^{\infty} w(\eta) e^{-j\omega\eta} d\eta, \quad W(-j\omega) = \int_{-\infty}^{\infty} w(\xi) e^{j\omega\xi} d\xi$$

Используя полученные выражения, можно записать

$$S_{yy}(\omega) = W(j\omega)W(-j\omega) \int_{-\infty}^{\infty} K_{xx}(\xi + \tau - \eta) e^{-j\omega(\xi + \tau - \eta)} d\tau.$$

Но поскольку справедливы соотношения

$$W(j\omega)W(-j\omega) = |W(j\omega)|^2;$$

$$\int_{-\infty}^{\infty} K_{xx}(\xi + \tau - \eta) e^{-j\omega(\xi + \tau - \eta)} d\tau = S_{xx}(\omega),$$

окончательно получаем

$$S_{yy}(\omega) = |W(j\omega)|^2 S_{xx}(\omega) = R^2(\omega) S_{xx}(\omega).$$

Следовательно, *спектральная плотность стационарного случайного сигнала на выходе линейного звена определяется произведением спектральной плотности стационарного случайного сигнала на входе звена на квадрат амплитудно-частотной характеристики этого звена.*

Из полученной формулы следует, что фазочастотная характеристика $\phi(\omega)$ звена, не входящая в выражение для $S_{yy}(\omega)$, не оказывает влияния на спектральную плотность выходного сигнала.

В частном случае прохождения случайного сигнала через идеальное дифференцирующее звено с передаточной функцией $W(s) = ks$ спектральная плотность сигнала на выходе равна

$$S_{yy}(\omega) = k^2 \omega^2 S_{xx}(\omega),$$

т. е. дифференцирующее звено ослабляет сигналы низких частот и усиливает сигналы высоких частот. Соответственно, спектральная плотность сигнала на выходе интегрирующего звена с передаточной функцией $W(s) = 1/Ts$ равна

$$S_{yy}(\omega) = (1/T^2 \omega^2) S_{xx}(\omega),$$

т. е. интегрирующее звено, наоборот, ослабляет сигналы высоких частот и усиливает сигналы низких частот.

Корреляционная функция сигнала на выходе звена, выражаемая через его спектральную плотность, в соответствии с формулой Винера-Хинчина (14.3) имеет вид

$$K_{yy}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |W(j\omega)|^2 S_{xx}(\omega) e^{j\omega\tau} d\omega$$

Эта формула удобнее для расчетов, чем (15.1), требующая двойного интегрирования.

Дисперсию сигнала на выходе при действии идеального белого шума со спектральной плотностью $S_{xx}(\omega) = c^2$ можно представить так:

$$D_y = \frac{c^2}{2\pi} \int_{-\infty}^{\infty} |W(j\omega)|^2 d\omega = \frac{c^2}{2\pi} \int_{-\infty}^{\infty} R^2(\omega) d\omega$$

В соответствии с ранее полученным выражением для взаимной спектральной плотности (см. п. 14.3.4) можно записать:

$$S_{yx}(j\omega) = \lim_{T \rightarrow \infty} \frac{1}{2T} Y_T(j\omega) X_T(-j\omega) = \lim_{T \rightarrow \infty} \frac{Y_T(j\omega)}{X_T(j\omega)} \lim_{T \rightarrow \infty} \frac{1}{2T} |X_T(j\omega)|^2.$$

Комплексный коэффициент передачи линейного звена может быть определен и таким образом:

$$W(j\omega) = \lim_{T \rightarrow \infty} \left. \frac{Y_T(s)}{X_T(s)} \right|_{s=j\omega},$$

в результате получаем

$$S_{yx}(j\omega) = W(j\omega) S_{xx}(\omega).$$

Таким образом, *взаимная спектральная плотность выходного и входного сигналов линейного звена равна произведению спектральной плотности входного сигнала на комплексный коэффициент передачи звена.*

15.2. ОБЪЕКТ УПРАВЛЕНИЯ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

Рассмотрим устойчивый линейный объект управления (рис. 15.2, а), к различным входам которого приложены стационарные случайные воздействия, причем $U(t)$ — управляющее воздействие, $F(t)$ — помеха. Структурная схема, отражающая преобразование объектом случайных воздействий, приведена на рис. 15.2, б, где $W_u(s)$, $W_o(s)$ — передаточные функции объекта по каналам управления и помехи соответственно. Возникает задача установления связи корреляционной функции и спектральной плотности сигнала на выходе с известными корреляционными функциями и спектральными плотностями сигналов на входах.

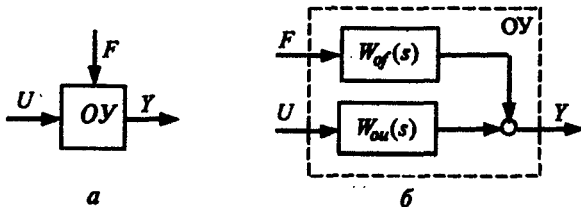


Рис. 15.2. Объект управления (а) и его структурная схема (б)

Поскольку для линейных систем справедлив принцип суперпозиции, то можно представить реализацию установившейся реакции объекта следующим образом:

$$y(t) = \int_{-\infty}^{\infty} w_{ou}(\xi) u(t-\xi) d\xi + \int_{-\infty}^{\infty} w_{of}(\xi) f(t-\xi) d\xi,$$

где $w_{ou}(\xi)$, $w_{of}(\xi)$ — весовые функции объекта по каналам управления и помехи; $u(t-\xi)$, $f(t-\xi)$ — реализации случайных воздействий.

Реализация установившейся реакции объекта, сдвинутой на время τ , записывается так:

$$y(t+\tau) = \int_{-\infty}^{\infty} w_{ou}(\eta) u(t+\tau-\eta) d\eta + \int_{-\infty}^{\infty} w_{of}(\eta) f(t+\tau-\eta) d\eta.$$

Тогда корреляционную функцию можно представить в виде

$$K_{yy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[\int_{-\infty}^{\infty} w_{ou}(\xi) u(t-\xi) d\xi + \int_{-\infty}^{\infty} w_{of}(\xi) f(t-\xi) d\xi \right] \times \\ \times \left[\int_{-\infty}^{\infty} w_{ou}(\eta) u(t+\tau-\eta) d\eta + \int_{-\infty}^{\infty} w_{of}(\eta) f(t+\tau-\eta) d\eta \right] dt.$$

Изменяя порядок интегрирования и осуществляя перемножение, получаем

$$K_{yy}(\tau) = \int_{-\infty}^{\infty} w_{ou}(\xi) d\xi \int_{-\infty}^{\infty} w_{ou}(\eta) K_{uu}(\xi+\tau-\eta) d\eta + \\ + \int_{-\infty}^{\infty} w_{ou}(\xi) d\xi \int_{-\infty}^{\infty} w_{of}(\eta) K_{uf}(\xi+\tau-\eta) d\eta + \\ + \int_{-\infty}^{\infty} w_{of}(\xi) d\xi \int_{-\infty}^{\infty} w_{ou}(\eta) K_{fu}(\xi+\tau-\eta) d\eta + \\ + \int_{-\infty}^{\infty} w_{of}(\xi) d\xi \int_{-\infty}^{\infty} w_{of}(\eta) K_{ff}(\xi+\tau-\eta) d\eta, \quad (15.3)$$

$$K_{uu}(\xi + \tau - \eta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T u(t - \xi) u(t + \tau - \eta) dt,$$

$$K_{ff}(\xi + \tau - \eta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t - \xi) f(t + \tau - \eta) dt,$$

$$K_{uf}(\xi + \tau - \eta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T u(t - \xi) f(t + \tau - \eta) dt,$$

$$K_{fu}(\xi + \tau - \eta) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t - \xi) u(t + \tau - \eta) dt.$$

Интегральное соотношение (15.3) определяет взаимосвязь между корреляционной функцией $K_{yy}(\tau)$ установившейся стационарной реакции объекта автокорреляционными функциями $K_{uu}(\tau)$, $K_{ff}(\tau)$ и взаимными корреляционными функциями $K_{uf}(\tau)$, $K_{fu}(\tau)$ стационарных случайных воздействий.

Спектральная плотность установившейся реакции объекта

$$S_{yy}(\omega) = |W_{ou}(j\omega)|^2 S_{uu}(\omega) + W_{ou}(-j\omega) W_{of}(j\omega) S_{uf}(j\omega) + \\ + W_{ou}(j\omega) W_{of}(-j\omega) S_{fu}(j\omega) + |W_{of}(j\omega)|^2 S_{ff}(\omega),$$

где $W_{ou}(-j\omega)$, $W_{of}(-j\omega)$ — частотные характеристики, комплексно-сопряженные с $W_{ou}(j\omega)$, $W_{of}(j\omega)$; $S_{uf}(j\omega)$, $S_{fu}(j\omega)$ — взаимные спектральные плотности воздействий.

В том случае, когда случайные воздействия $U(t)$ и $F(t)$ вероятностно независимы (не коррелированы), их взаимные спектральные плотности

$$S_{uf}(j\omega) = S_{fu}(j\omega) \equiv 0,$$

тогда

$$S_{yy}(\omega) = |W_{ou}(j\omega)|^2 S_{uu}(\omega) + |W_{of}(j\omega)|^2 S_{ff}(\omega).$$

15.3. ЗАМКНУТАЯ СИСТЕМА ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

Перейдем к рассмотрению замкнутой устойчивой системы управления с единичной обратной связью при случайных воздействиях (рис. 15.3). Цель управления такой системы состоит в воспроизведении полезного задающего воздействия $G(t)$ и подавлении возмущающего воздействия (помехи) $F(t)$, приложенного в любом месте объекта управления.

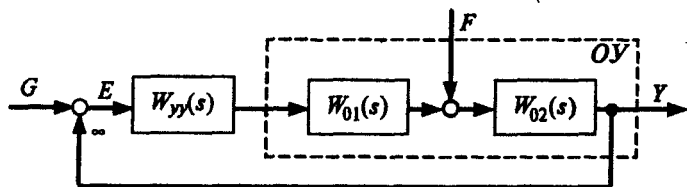


Рис. 15.3. Структурная схема замкнутой системы при случайных воздействиях

В общем случае воздействия можно представить в виде сумм:

$$G(t) = m_g(t) + \dot{G}(t); \quad F(t) = m_f(t) + \dot{F}(t),$$

где $m_g(t)$, $m_f(t)$ — эквивалентные регулярные составляющие, включающие математические ожидания соответствующих случайных воздействий и их регулярные (детерминированные) компоненты; $\dot{G}(t)$, $\dot{F}(t)$ — центрированные случайные составляющие воздействий.

Исследование точности воспроизведения задания $G(t)$ на выходе можно свести к рассмотрению ошибки системы

$$E(t) = G(t) - Y(t),$$

которую, в свою очередь, можно представить как

$$E(t) = m_e(t) + \dot{E}(t),$$

где $m_e(t)$ — математическое ожидание ошибки; $\dot{E}(t)$ — центрированная случайная составляющая ошибки.

Для линейных систем, в которых справедлив принцип суперпозиции, возможно нахождение составляющих ошибки $m_e(t)$ и $\dot{E}(t)$ от действия каждого из сигналов $G(t)$ и $F(t)$ в отдельности.

Регулярная составляющая ошибки $m_e(t)$ представляет собой реакцию на внешние регулярные составляющие воздействий $m_g(t)$ и $m_f(t)$. Ее изображение по Лапласу определяется так:

$$m_e(s) = \Phi_{eg}(s)m_g(s) + \Phi_{ef}(s)m_f(s).$$

В случае, когда регулярные составляющие воздействий постоянные, а случайные составляющие стационарны, математическое ожидание ошибки определяется выражением

$$m_e = \Phi_{eg}(0)m_g + \Phi_{ef}(0)m_f.$$

Случайная составляющая ошибки $\dot{E}(t)$ является реакцией системы на центрированные случайные составляющие воздействий $\dot{G}(t)$ и $\dot{F}(t)$. Вероятностные характеристики воздействий $\dot{G}(t)$ и $\dot{F}(t)$ обычно зада-

ются либо корреляционными функциями $K_{gg}(\tau)$ и $K_{ff}(\tau)$, либо спектральными плотностями $S_{gg}(\omega)$ и $S_{ff}(\omega)$. При коррелированности воздействий, кроме того, задается взаимная корреляционная функция $K_{gf}(\tau)$ или взаимная спектральная плотность $S_{gf}(j\omega)$. По заданным вероятностным характеристикам воздействий вычисляется дисперсия для сигнала ошибки D_e . На практике D_e чаще всего определяется по спектральным плотностям входных воздействий:

$$D_e = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{ee}(\omega) d\omega = \frac{1}{\pi} \int_0^{\infty} S_{ee}(\omega) d\omega,$$

где

$$S_{ee}(\omega) = |\Phi_{eg}(j\omega)|^2 S_{gg}(\omega) + \Phi_{eg}(-j\omega) \Phi_{ef}(j\omega) S_{gf}(j\omega) + \\ + \Phi_{eg}(j\omega) \Phi_{ef}(-j\omega) S_{fg}(j\omega) + |\Phi_{ef}(j\omega)|^2 S_{ff}(\omega),$$

$\Phi_{eg}(j\omega)$, $\Phi_{ef}(j\omega)$ — комплексные коэффициенты передач замкнутой системы по ошибке. Отсюда дисперсия ошибки будет складываться из отдельных составляющих:

$$D_e = D_e^g + D_e^{gf} + D_e^{fg} + D_e^f,$$

причем верхние индексы отвечают воздействиям, от которых зависит соответствующая составляющая дисперсии.

Если задающее и возмущающее воздействия не коррелированы, то их взаимные спектральные плотности равны нулю. Тогда спектральная плотность ошибки

$$S_{ee}(\omega) = |\Phi_{eg}(j\omega)|^2 S_{gg}(\omega) + |\Phi_{ef}(j\omega)|^2 S_{ff}(\omega).$$

Выражение для дисперсии ошибки упрощается:

$$D_e = D_e^g + D_e^f.$$

В частном случае, когда возмущение приложено непосредственно к выходу объекта ($W_{02}(s) = 1$), задающий и возмущающий сигналы не коррелированы, спектральная плотность ошибки

$$S_{ee}(\omega) = \left| \frac{1}{1+W(j\omega)} \right|^2 [S_{gg}(\omega) + S_{ff}(\omega)],$$

где $W(j\omega) = W_y(j\omega)W_{01}(j\omega)$.

Дисперсия и математическое ожидание ошибки определяют: среднее значение квадрата ошибки

$$\overline{e^2(t)} = D_e + m_e^2(t);$$

среднюю квадратическую ошибку

$$e_{\text{ско}}(t) = \sqrt{e^2(t)} = \sqrt{D_e + m_e^2(t)}.$$

При постоянных регулярных составляющих и стационарных случайных составляющих воздействий получаем:

$$\overline{e^2} = D_e + m_e^2; \quad e_{\text{ско}} = \sqrt{D_e + m_e^2}.$$

В частности, когда регулярные составляющие воздействий отсутствуют, т. е. $m_e = 0$, среднеквадратическая ошибка

$$e_{\text{ско}} = \sqrt{D_e} = \sigma_e,$$

где σ_e — среднеквадратическое отклонение ошибки.

15.4. СПОСОБЫ ВЫЧИСЛЕНИЯ ДИСПЕРСИИ СЛУЧАЙНОГО СИГНАЛА

15.4.1. Аналитический способ

Аналитическое определение значения дисперсии случайного сигнала $Y(t)$ системы основано на применении формулы

$$D_y = \frac{1}{2\pi} \int_{-\infty}^{\infty} |W_{yx}(j\omega)|^2 S_{xx}(\omega) d\omega, \quad (15.4)$$

где $W_{yx}(j\omega)$ — комплексный коэффициент передачи между переменными x и y , $S_{xx}(\omega)$ — спектральная плотность случайного сигнала $X(t)$.

В связи с тем, что передаточная функция $W_{yx}(s)$ представляет собой отношение двух полиномов $B_{yx}(s)$ и $A_{yx}(s)$, квадрат модуля комплексного коэффициента передачи является функцией четных степеней ω , т. е.

$$|W_{yx}(j\omega)|^2 = W_{yx}(\omega^2).$$

Справедливость этого равенства следует из того, что квадрат модуля любого полинома $G(j\omega)$ степени r может быть представлен выражением

$$\begin{aligned} |G(j\omega)|^2 &= |g_r(j\omega)^r + \dots + g_1(j\omega) + g_0|^2 = |H[(j\omega)^2] + j\omega L[(j\omega)^2]|^2 = \\ &= \{H[(j\omega)^2]\}^2 + \omega^2 \{L[(j\omega)^2]\}^2 = G(\omega^2), \end{aligned}$$

где $H[(j\omega)^2] = g_0 + g_2(j\omega)^2 + \dots$, $L[(j\omega)^2] = g_1 + g_3(j\omega)^2 + \dots$

Будем считать, что спектральная плотность $S_{xx}(\omega)$ представляет собой неотрицательную, вещественную и четную дробно-рациональную функцию, полиномы которой содержат члены только четных степеней ω , т. е.

$$S_{xx}(\omega) = \frac{P(\omega^2)}{Q(\omega^2)} = \frac{p_m \omega^{2m} + p_{m-1} \omega^{2m-2} + \dots + p_0}{q_n \omega^{2n} + q_{n-1} \omega^{2n-2} + \dots + q_0}, \quad m < n,$$

где p_r, q_l — действительные коэффициенты полиномов P и Q .

Степень $2m$ полинома $P(\omega^2)$ числителя всегда меньше степени $2n$ полинома $Q(\omega^2)$ знаменателя, так как только в этом случае интеграл от спектральной плотности, определяющий дисперсию случайного сигнала $X(t)$, сходится и дисперсия D_x будет конечной величиной. Кроме того, для сходимости интеграла от спектральной плотности необходимо, чтобы ее знаменатель не обращался в нуль ни при каком значении частоты ω .

Корни уравнений $P(\omega^2) = 0$, $Q(\omega^2) = 0$ с действительными значениями коэффициентов либо действительны, либо попарно комплексно-сопряженные. Если полином является четной функцией, то каждому корню ω_k соответствует корень $-\omega_k$.

Например, полином

$$F(\omega^2) = f_n \omega^{2n} + f_{n-1} \omega^{2n-2} + \dots + f_0$$

при обозначении $p = \omega^2$, можно представить как

$$F(p) = f_n p^n + f_{n-1} p^{n-1} + \dots + f_0.$$

Раскладывая его на элементарные сомножители, получим

$$F(p) = f_n \prod_{k=1}^r (p - p_k) \prod_{k=r+1}^n [(p - \alpha_k)^2 + \beta_k^2] = 0.$$

При этом корню $p_k = \omega_k^2 \geq 0$ будет соответствовать пара симметричных действительных корней $\omega_{k1} = \sqrt{p_k}$, $\omega_{k2} = -\sqrt{p_k}$; корню $p_k = \omega_k^2 < 0$ — пара симметричных мнимых корней $\omega_{k1} = j\sqrt{p_k}$, $\omega_{k2} = -j\sqrt{p_k}$. Пары комплексно-сопряженных корней $p_{k1} = \alpha_k + j\beta_k$, $p_{k2} = \alpha_k - j\beta_k$ будут соответствовать четыре комплексных корня:

$$\omega_{k1} = u_k + jv_k, \quad \omega_{k1}^2 = u_k^2 - v_k^2 + j2u_kv_k;$$

$$\omega_{k2} = -u_k - jv_k, \quad \omega_{k2}^2 = u_k^2 - v_k^2 + j2u_kv_k;$$

$$\omega_{k3} = u_k - jv_k, \quad \omega_{k3}^2 = u_k^2 - v_k^2 - j2u_kv_k;$$

$$\omega_{k4} = -u_k + jv_k, \quad \omega_{k4}^2 = u_k^2 - v_k^2 - j2u_kv_k.$$

Полином $Q(\omega^2)$ знаменателя не может содержать действительных корней, так как иначе $S_{xx}(\omega)$ будет иметь полюс на положительной действительной полуоси и интеграл от этой функции не будет конечным.

Возможное расположение корней полиномов $P(\omega^2)$ и $Q(\omega^2)$ на плоскости комплексной переменной $\omega = u + jv$ представлено на рис. 15.4.

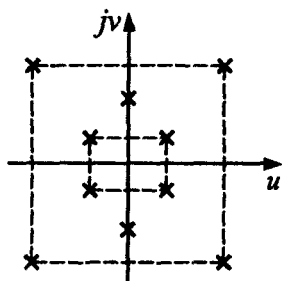


Рис. 15.4. Расположение корней в плоскости комплексной переменной ω

Корни располагаются симметрично относительно действительной и мнимой осей.

Таким образом, подынтегральная функция выражения (15.4) для вычисления дисперсии D_y является функцией только четных степеней ω :

$$S_{xx}(\omega) |W_{yx}(j\omega)|^2 = \frac{B(\omega^2)}{A(\omega^2)}.$$

В случае простых и вещественных полюсов $p_i = \omega_i^2$ полученное выражение может быть представлено суммой простейших дробей:

$$\frac{B(\omega^2)}{A(\omega^2)} = \sum_{i=1}^n \frac{c_i}{\omega^2 + \omega_i^2},$$

где n — число полюсов $p_i = \omega_i^2$.

В результате интегрирования этого выражения находим

$$D_y = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sum_{i=1}^n \frac{c_i}{\omega^2 + \omega_i^2} d\omega = \frac{1}{2\pi} \sum_{i=1}^n \frac{c_i}{\omega_i} \operatorname{arctg} \frac{\omega}{\omega_i} \Big|_{-\infty}^{\infty} = \frac{1}{2} \sum_{i=1}^n \frac{c_i}{\omega_i}.$$

Следовательно, для вычисления дисперсии D_y необходимо найти корни уравнения $A(\omega^2) = 0$ и определить коэффициенты c_i . Эти вычисления существенно усложняются, когда корни уравнения комплексные, что имеет место в большинстве случаев на практике.

В общем случае для вычисления дисперсии D_y по формуле (15.4) разработан аналитический способ, основанный на представлении подынтегральной функции в следующем виде:

$$\begin{aligned} D_y = J_n &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{B(\omega^2)}{A(\omega^2)} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{|B(j\omega)|^2}{A(j\omega)A(-j\omega)} d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{b_{n-1}(j\omega)^{2(n-1)} + b_{n-2}(j\omega)^{2(n-2)} + \dots + b_1(j\omega)^2 + b_0}{A(j\omega)A(-j\omega)} d\omega, \end{aligned} \quad (15.5)$$

где $A(j\omega) = a_n(j\omega)^n + a_{n-1}(j\omega)^{n-1} + \dots + a_0$
 $A(-j\omega) = a_n(-j\omega)^n + a_{n-1}(-j\omega)^{n-1} + \dots + a_0.$

Знаменатель подынтегрального выражения допускает такое представление вследствие справедливости соотношений:

$$\begin{aligned}
A(\omega^2) &= |A(j\omega)|^2 = \{H[(j\omega)^2]\}^2 + \omega^2 \{L[(j\omega)^2]\}^2 = \\
&= \{H[(j\omega)^2]\}^2 - (j\omega)^2 \{L[(j\omega)^2]\}^2 = \\
&= \{H[(j\omega)^2] + j\omega L[(j\omega)^2]\} \{H[(j\omega)^2] - j\omega L[(j\omega)^2]\} = A(j\omega)A(-j\omega).
\end{aligned}$$

Полиномы $H[(j\omega)^2]$ и $L[(j\omega)^2]$ содержат члены только с четными степенями $(j\omega)$. Замена $(j\omega)$ на $(-j\omega)$ не меняет их вида, так как $(j\omega)^2 = (-j\omega)^2$. Тогда второй множитель

$$A(-j\omega) = H[(-j\omega)^2] + (-j\omega)L[(-j\omega)^2] = a_n(-j\omega)^n + a_{n-1}(-j\omega)^{n-1} + \dots + a_0.$$

Корни полинома $A(j\omega)$ расположены в верхней полуплоскости комплексной плоскости, корни полинома $A(-j\omega)$ — в нижней полуплоскости.

Справедливость последнего утверждения вытекает из следующих рассуждений. Разложив полином $A(\omega^2)$ на элементарные множители, выберем те, которые отвечают корням, расположенным в верхней полуплоскости комплексной плоскости ω . Полученное разложение умножим на коэффициент $j^n \sqrt{a_n}$. В результате получим полином $A(j\omega)$ степени n относительно ω , корни которого расположены в верхней полуплоскости. Коэффициенты полинома $A(j\omega)$ относительно аргумента $j\omega$ положительны.

Действительно, полином $A(j\omega)$ представляет собой произведение положительного числа $\sqrt{a_n}$ на множители вида $j(\omega - u - jv)j(\omega + u - jv)$, соответствующие «верхним» комплексным корням, и множители вида $j(\omega - jv)$, соответствующие «верхним» мнимым корням. Эти множители можно преобразовать к виду

$$j(\omega - u - jv)j(\omega + u - jv) = (j\omega)^2 + 2v(j\omega) + u^2 + v^2 \quad \text{и} \quad j(\omega - jv) = j\omega + v,$$

вследствие чего произведение любого числа таких множителей представляет собой полином относительно $j\omega$, все коэффициенты которого положительны.

Оставшиеся множители в разложении полинома $A(\omega^2)$, соответствующие корням, расположенным в нижней полуплоскости, умноженные на $(-j)^n \sqrt{a_n}$, образуют полином, который получается из $A(j\omega)$ изменением знака у $j\omega$, т. е. полином $A(-j\omega)$.

Действительно, множители полинома $A(\omega^2)$, образованные комплексными корнями вида $u - jv$ и $-u - jv$, и мнимыми корнями вида $-jv$, расположенными в нижней полуплоскости, преобразовываются к выражениям

$$-j(\omega - u + jv)(-j)(\omega + u + jv) = (-j\omega)^2 + 2v(-j\omega) + u^2 + v^2;$$

$$-j(\omega + jv) = -j\omega + v,$$

что и доказывает высказанное утверждение.

Таким образом, выбрав в разложении полинома $A(\omega^2)$ множители, соответствующие корням, расположенным в верхней полуплоскости, и множители, соответствующие корням, расположенным в нижней полуплоскости, и умножив их на $j^n \sqrt{a_n}$ и $(-j)^n \sqrt{a_n}$ соответственно, можем получить

$$A(\omega^2) = A(j\omega)A(-j\omega),$$

где $A(j\omega)$ — полином относительно аргумента $j\omega$ с положительными коэффициентами.

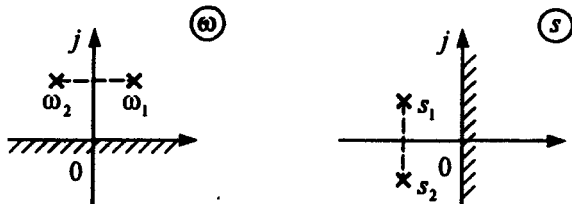


Рис. 15.5. Отображение комплексной плоскости ω в комплексную плоскость s

Аналогичным образом можно представить полином числителя

$$B(\omega^2) = B(j\omega)B(-j\omega),$$

где $B(j\omega)$ — полином относительно аргумента $j\omega$ с положительными коэффициентами.

Следует заметить, что умножение комплексного числа на мнимую единицу j представляет собой поворот вектора, отображающего это число, на угол $\pi/2$ против часовой стрелки. Следовательно, все располагаемые в верхней полуплоскости комплексной ω корни полиномов $A(j\omega)$ и $B(j\omega)$, рассматриваемых как функции ω , будут располагаться в левой полуплоскости комплексной плоскости s и являться корнями полиномов $A(s)$ и $B(s)$, рассматриваемых как функции s (рис. 15.5).

Интегралы представленного вида (15.5) для устойчивой системы при любом n вычисляются по формуле

$$J_n = \frac{T_n}{2a_n\Delta_n},$$

где Δ_n — старший определитель матрицы Гурвица (см. п. 3.3.2), составленный из коэффициентов полинома $A(j\omega)$, т. е.

$$\Delta_n = \begin{vmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \dots & 0 \\ a_n & a_{n-2} & a_{n-4} & \dots & 0 \\ 0 & a_{n-1} & a_{n-3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_0 \end{vmatrix},$$

T_n — определитель, совпадающий с Δ_n , но в котором первая строка заменена коэффициентами $b_{n-1}, b_{n-2}, \dots, b_0$ полинома $B(j\omega)$, т. е.

$$T_n = \begin{vmatrix} b_{n-1} & b_{n-2} & b_{n-3} & \dots & b_0 \\ a_n & a_{n-2} & a_{n-4} & \dots & 0 \\ 0 & a_{n-1} & a_{n-3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_0 \end{vmatrix}.$$

В частности, для значений $n = 1 + 4$ выражения для J_n приведены в табл. 15.1.

Таблица 15.1

Степень полинома n	Выражение для интеграла J_n
1	$\frac{b_0}{2a_1a_0}$
2	$\frac{-b_1 + b_0 a_2/a_0}{2a_2a_1}$
3	$\frac{-a_1b_2 + a_3b_1 - a_2a_2b_0/a_0}{2a_1(a_3a_0 - a_2a_1)}$
4	$\frac{b_2(-a_3a_0 + a_2a_1) - a_4a_1b_2 + a_4a_2b_1 + a_4b_0(a_4a_1 - a_3a_2)/a_0}{2a_4(a_4a_1^2 + a_3^2a_0 - a_3a_2a_1)}$

Знаменатели формул для J_n содержат множителями определители Гурвица Δ_{n-1} . В том случае, если в системе имеет место колебательная граница устойчивости, т. е. имеется пара мнимых корней характеристического уравнения, то $\Delta_{n-1} = 0$ и дисперсии сигналов в системе стремятся к бесконечности.

Пример. Спектральная плотность случайного сигнала $X(t)$ на входе системы с передаточной функцией

$$W_{yx}(s) = \frac{1}{s^2 + s + 1}$$

имеет вид

$$S_{xx}(\omega) = \frac{4}{\omega^2 + 1}.$$

Требуется определить дисперсию сигнала $Y(t)$ на выходе системы.

$$S_{yy}(\omega) = \frac{4}{|(j\omega)^2 + j\omega + 1|^2 (\omega^2 + 1)}.$$

Для определения дисперсии D_y в соответствии с (15.5) подынтегральная функция приводится к виду

$$\begin{aligned} S_{yy}(\omega) &= \frac{4}{[(j\omega)^2 + j\omega + 1](j\omega + 1)[(-j\omega)^2 + (-j\omega) + 1](-j\omega + 1)} = \\ &= \frac{4}{[(j\omega)^3 + 2(j\omega)^2 + 2(j\omega) + 1][(-j\omega)^3 + 2(-j\omega)^2 + 2(-j\omega) + 1]} \end{aligned}$$

Теперь можно воспользоваться формулой для $n = 3$ из табл. 15.1:

$$D_y = J_3 = \frac{-a_1 b_2 + a_3 b_1 - a_3 a_2 b_0 / a_0}{2a_3(a_3 a_0 - a_2 a_1)},$$

где $b_0 = 4$, $b_2 = b_1 = 0$; $a_0 = 1$, $a_1 = 2$, $a_2 = 2$, $a_3 = 1$. При подстановке значений коэффициентов в D_y окончательно получаем

$$D_y = 4/3.$$

15.4.2. Графоаналитический способ

В инженерной практике часто используется графоаналитический способ определения дисперсии случайного сигнала. Поясним сущность этого способа применительно к вычислению дисперсии ошибки D_e , вызванной действием на замкнутую систему управления некоррелированных случайных сигналов $G(t)$ и $F(t)$ (см. рис. 15.3).

Сначала строятся графики амплитудно-частотной характеристики замкнутой системы $R_{ef}(\omega) = |\Phi_{ef}(j\omega)|$, связывающей ошибку $E(t)$ с возмущением $F(t)$, и спектральной плотности возмущения $S_{ff}(\omega)$ (рис. 15.6).

Затем строится график функции $R_{ef}^2(\omega)$. Перемножая $R_{ef}^2(\omega)$ и $S_{ff}(\omega)$ при одних и тех же значениях частоты ω , получаем график спектральной плотности $S_{ee}^f(\omega)$. В результате таких построений площадь области, ограниченной кривой $S_{ee}^f(\omega)$, осями абсцисс и ординат, деленной на множитель π , дает значение составляющей дисперсии ошибки по возмущению

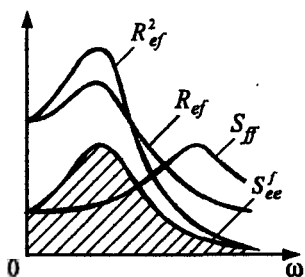


Рис. 15.6. Графоаналитическое определение дисперсии ошибки

$$D_{\epsilon}^f = \frac{1}{\pi} \int_0^{\infty} S_{\epsilon\epsilon}^f(\omega) d\omega.$$

Аналогичным образом можно найти составляющую дисперсии ошибки D_{ϵ}^g по заданию. Суммируя эти составляющие, определяется дисперсия ошибки D_{ϵ} .

Графоаналитический способ определения дисперсии особенно удобен для применения, когда характеристики $R_{\epsilon f}(\omega)$, $R_{\epsilon g}(\omega)$ и $S_{ff}(\omega)$, $S_{gg}(\omega)$ получены экспериментально и задаются в виде графиков.

15.5. ПРИМЕНЕНИЕ ИДЕАЛЬНОГО БЕЛОГО ШУМА В КАЧЕСТВЕ МОДЕЛИ ВНЕШНЕЙ СРЕДЫ

Выбор моделей внешней среды производится, с одной стороны, с учетом относительной простоты описания воздействий, а с другой — наибольшего соответствия реальности. В практике расчетов систем управления широко используется модель воздействия в виде идеального белого шума.

Пусть на входах системы (см. рис. 15.3) с передаточной функцией $W(s) = W_{yy}(s)W_{01}(s) = k/s$, ($W_{02}(s) = 1$) в прямой цепи действуют стационарные случайные сигналы $G(t)$ и $F(t)$, представленные в виде идеального белого шума $N(t)$ со спектральной плотностью $S_{nn}(\omega) = c^2$.

Дисперсия реакции на выходе системы от действия $G(t)$ имеет вид

$$D_y^g = \frac{1}{2\pi} \int_{-\infty}^{\infty} c^2 \left| \frac{k}{j\omega + k} \right|^2 d\omega = \frac{c^2 k^2}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega}{(j\omega + k)(-j\omega + k)} = c^2 k^2 J_1 = c^2 k^2 \frac{b_0}{2a_1 a_0},$$

причем $b_0 = 1$; $a_1 = 1$; $a_0 = k$. Следовательно,

$$D_y^g = c^2 k/2.$$

Дисперсия сигнала ошибки от действия $G(t)$, а также дисперсия реакции на выходе системы от действия $F(t)$ определяется так:

$$D_{\epsilon}^g = D_y^f = \frac{1}{2\pi} \int_{-\infty}^{\infty} c^2 \left| \frac{j\omega}{j\omega + k} \right|^2 d\omega = \frac{c^2}{2\pi} \int_{-\infty}^{\infty} \frac{\omega^2 d\omega}{\omega^2 + k^2} = \frac{c^2}{2\pi} \int_{-\infty}^{\infty} d\omega - D_y^g.$$

Первое слагаемое в выражении для дисперсии принимает бесконечное значение, что не может соответствовать реальному случаю, так как в действительности дисперсия сигналов должна иметь ограниченное значение. Полученный результат объясняется тем, что выражения для квадратов модулей комплексных коэффициентов передач $|\Phi_{\epsilon g}(j\omega)|^2$ и $|\Phi_{yy}(j\omega)|^2$

содержат степени полиномов числителя, равные степеням полиномов знаменателя.

Передаточная функция системы

$$\Phi_{yg}(s) = \frac{k}{s+k}$$

соответствует *фильтру низких частот*, который пропускает сигналы низких частот и подавляет сигналы высоких частот.

Передаточные функции системы по ошибке и по возмущению

$$\Phi_{eg}(s) = \Phi_{yf}(s) = \frac{s}{s+k}$$

отвечают *фильтру высоких частот*, подавляющего сигналы низких и пропускающего сигналы высоких частот.

Квадраты модулей комплексных коэффициентов передач фильтров низких и высоких частот соответственно имеют вид:

$$|\Phi_{yg}(j\omega)|^2 = R_{\text{фнч}}^2(\omega) = \frac{k^2}{\omega^2 + k^2};$$

$$|\Phi_{eg}(j\omega)|^2 = |\Phi_{yf}(j\omega)|^2 = R_{\text{фвч}}^2(\omega) = \frac{\omega^2}{\omega^2 + k^2}.$$

Их графики приведены на рис. 15.7.

Из рассмотренного примера следует, что использование модели воздействия в виде идеального белого шума для расчета дисперсии сигналов системы возможно лишь тогда, когда передаточная функция от воздействия до исследуемой переменной системы отвечает фильтру низких частот.

В том случае, когда передаточная функция между входом и исследуемой переменной системы соответствует фильтру высоких частот, следует переходить к модели воздействия с *ограниченной полосой частотного спектра* $\bar{S}_m(\omega)$ (см. рис. 15.7), например

$$\bar{S}_{nn}(\omega) = \frac{2D\mu}{\omega^2 + \mu^2}.$$

Этот вывод распространяется и на системы высокого порядка. Действительно, в том случае, если передаточная функция разомкнутой системы

$$W(s) = \frac{k(\tau_1 s + 1) \dots (\tau_m s + 1)}{s(T_1 s + 1) \dots (T_n s + 1)}, \quad m \leq n,$$

то передаточные функции замкнутой системы по ошибке и по возмущению

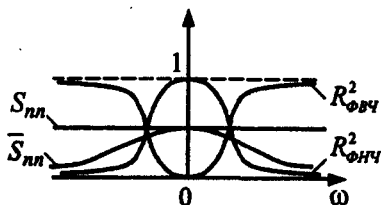


Рис. 15.7. Графики квадратов модулей комплексных коэффициентов передач фильтров низких и высоких частот

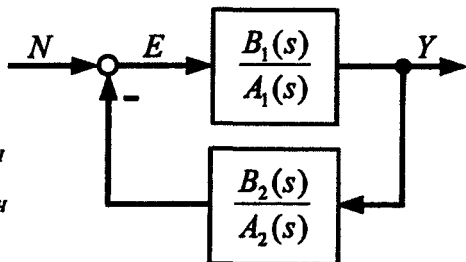


Рис. 15.8. Применение идеального белого шума в системе произвольного порядка

$$\Phi_{eg}(s) = \Phi_{yf}(s) = \frac{s(T_1 s + 1) \dots (T_n s + 1)}{s(T_1 s + 1) \dots (T_n s + 1) + k(\tau_1 s + 1) \dots (\tau_m s + 1)}.$$

Вследствие одинаковости степеней полиномов числителя и знаменателя, передаточные функции $\Phi_{eg}(s)$ и $\Phi_{yf}(s)$ соответствуют фильтру высоких частот и поэтому модель воздействия в виде идеального белого шума на входах системы для расчета дисперсий D_x^g и D_y^f неприемлема.

Если передаточная функция прямой цепи системы (рис. 15.8)

$$W(s) = B_1(s)/A_1(s),$$

передаточная функция обратной связи

$$W_{oc}(s) = B_2(s)/A_2(s),$$

причем степени полиномов $B_2(s)$ и $A_2(s)$ отвечают условию физической осуществимости ($\deg B_2 \leq \deg A_2$), то передаточная функция замкнутой системы

$$\Phi(s) = \frac{B_1(s)A_2(s)}{A_1(s)A_2(s) + B_1(s)B_2(s)} = \frac{B(s)}{A(s)}.$$

В случае, когда $\deg B_1 = \deg A_1$, выполняется $\deg B = \deg A$ и идеальный белый шум $N(t)$ не может использоваться в качестве модели внешней среды при определении дисперсии D_x^g .

Получить необходимую модель внешней среды можно с помощью формирующего фильтра.

Формирующий фильтр — это динамическая система, на выходе которой формируется требуемый процесс, соответствующий решению определенного дифференциального уравнения. На вход такой динамической системы подается типовой сигнал. Для генерирования детерминированного сигнала подаются δ -функция или единичная ступенчатая

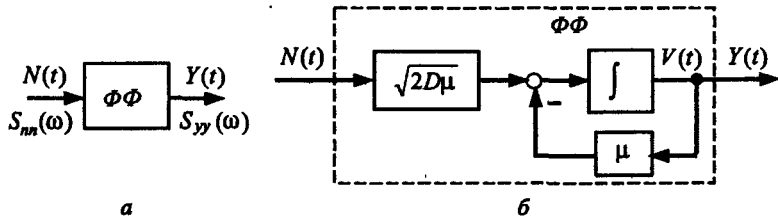


Рис. 15.9. Формирующий фильтр (а) и его структурная схема (б)

функция (см. 2.14). Для генерирования случайного сигнала с заданными вероятностными характеристиками подается идеальный белый шум.

На рис. 15.9, а изображен формирующий фильтр (ΦΦ), возбуждаемый идеальным белым шумом со спектральной плотностью $S_{nn}(\omega) = 1$. Требуется определить передаточную функцию $W_{\Phi\Phi}(s)$ фильтра, обеспечивающего на выходе случайный сигнал с заданной спектральной плотностью, например $S_{yy}(\omega) = 2D\mu/(\omega^2 + \mu^2)$.

Условие преобразования фильтром спектральной плотности имеет вид

$$S_{yy}(\omega) = |W_{\Phi\Phi}(j\omega)|^2 S_{nn}(\omega).$$

Подставляя заданные выражения, получаем

$$\frac{2D\mu}{\omega^2 + \mu^2} = \frac{2D\mu}{(j\omega + \mu)(-j\omega + \mu)} = |W_{\Phi\Phi}(j\omega)|^2 \cdot 1.$$

Учитывая, что комплексный коэффициент передачи фильтра должен отвечать минимально-фазовому звену, можно записать

$$W_{\Phi\Phi}(j\omega) = \frac{\sqrt{2D\mu}}{j\omega + \mu}.$$

При замене $j\omega$ на s находим передаточную функцию фильтра

$$W_{\Phi\Phi}(s) = \frac{\sqrt{2D\mu}}{s + \mu} = \frac{k_{\Phi\Phi}}{T_{\Phi\Phi}s + 1},$$

где $k_{\Phi\Phi} = \sqrt{2D}/\mu$, $T_{\Phi\Phi} = \mu^{-1}$.

Описание формирующего фильтра в форме пространства состояний (в форме Коши) имеет следующий вид:

$$\begin{aligned} \dot{V}(t) &= -\mu V(t) + \sqrt{2D\mu} N(t); \\ Y(t) &= V(t). \end{aligned}$$

На рис. 15.9, б приведена структурная схема реализации формирующего фильтра, обеспечивающего на выходе заданную спектральную плотность $S_{yy}(\omega)$.

15.6. ВЫЧИСЛЕНИЕ ДИСПЕРСИИ ОШИБКИ В СИСТЕМАХ РАЗЛИЧНОГО ПОРЯДКА

15.6.1. Оценка точности преобразования стационарных случайных сигналов в установившихся режимах

В качестве оценки точности преобразования полезного стационарного случайного сигнала в установившихся режимах часто используется дисперсия (или среднеквадратическое отклонение) ошибки системы, характеризующая степень разброса отклонений случайного сигнала относительно его математического ожидания. Являясь по определению квадратичной оценкой, величина дисперсии ошибки значительно возрастает, т. е. ее «вес» усиливается при больших отклонениях случайного сигнала.

Анализ точности преобразования случайных сигналов в системах управления различного порядка в зависимости от параметров передаточных функций системы и характеристик приложенных воздействий производится по структурной схеме расчетного вида, изображенной на рис. 15.10 [35, 69]. Вследствие принципа суперпозиции анализ может быть осуществлен раздельно для сигналов $M(t)$, $N(t)$ и $F(t)$. Полезный сигнал $M(t)$, поступающий в аддитивной смеси с помехой $N(t)$, необходимо без искажений воспроизводить на выходе, помеха $N(t)$ на выходе должна подавляться системой, сигнал возмущения $F(t)$, приведенный к выходу объекта управления, также подлежит ослаблению на выходе системы.

В результате прохождения входных сигналов через систему установившаяся реакция на выходе может быть представлена так:

$$Y(t) = Y_m(t) + Y_f(t) + Y_n(t),$$

где $Y_m(t)$ — полезный сигнал на выходе; $Y_f(t)$ — возмущение на выходе; $Y_n(t)$ — помеха на выходе.

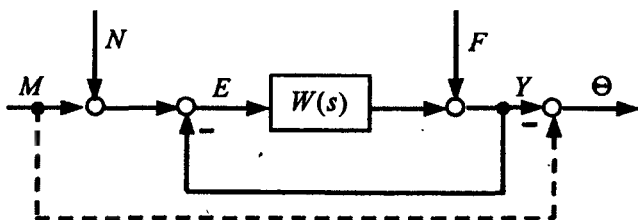


Рис. 15.10. Структурная схема расчетного вида

Установившаяся ошибка преобразования системой полезного случайного сигнала $M(t)$ в выходной сигнал $Y(t)$ определяется выражением

$$\Theta(t) = M(t) - Y(t) = E(t) - Y_f(t) - Y_n(t),$$

где $E(t) = M(t) - Y_m(t)$ — ошибка обработки системой полезного сигнала.

При отсутствии помехи $N(t)$ и возмущения $F(t)$ ошибка системы $\Theta(t) = E(t)$.

Дисперсия ошибки D_Θ , вызванная приложенными к системе вероятностно независимыми воздействиями, складывается из следующих составляющих:

$$D_\Theta = D_e^m + D_y^f + D_y^n,$$

где D_e^m — дисперсия ошибки системы при наличии только полезного сигнала $M(t)$; D_y^f — дисперсия возмущения на выходе системы; D_y^n — дисперсия помехи на выходе системы.

Составляющие дисперсии ошибки:

$$D_e^m = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| 1 - \frac{W(j\omega)}{1 + W(j\omega)} \right|^2 S_{mm}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi_{em}(j\omega)|^2 S_{mm}(\omega) d\omega;$$

$$D_y^f = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{1}{1 + W(j\omega)} \right|^2 S_{ff}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi_{yf}(j\omega)|^2 S_{ff}(\omega) d\omega;$$

$$D_y^n = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{W(j\omega)}{1 + W(j\omega)} \right|^2 S_{nn}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\Phi(j\omega)|^2 S_{nn}(\omega) d\omega,$$

где $S_{mm}(\omega)$, $S_{ff}(\omega)$, $S_{nn}(\omega)$ — спектральные плотности воздействий.

Таким образом, для определения точности системы управления в установившемся режиме, необходимо вычислить значения составляющих дисперсии ошибки при известных спектральных плотностях входных сигналов и комплексных коэффициентов передач системы.

Будем полагать, что сигналы $M(t)$ и $F(t)$ характеризуются спектральной плотностью $S_{mm}(\omega) = S_{ff}(\omega) = 2D\mu/(\omega^2 + \mu^2)$, а помеха $N(t)$ имеет спектральную плотность типа идеального белого шума, т. е. $S_{nn}(\omega) = c^2$.

Для систем высокого порядка выражения для дисперсии сигналов имеют сложный вид. Однако результаты качественного анализа, проводимого ниже для систем низкого порядка, во многом остаются справедливыми и для систем высокого порядка.

15.6.2. Вычисление дисперсии ошибки в системах первого порядка

Пусть передаточная функция разомкнутой системы соответствует интегрирующему звену

$$W(s) = \frac{k}{s},$$

помеха на входе отсутствует. Тогда передаточная функция замкнутой системы

$$\Phi(s) = \frac{k}{s+k},$$

передаточные функции замкнутой системы по ошибке $E(t)$ и по возмущению $F(t)$:

$$\Phi_{em}(s) = \Phi_{yf}(s) = \frac{s}{s+k}.$$

Квадраты модулей комплексных коэффициентов передачи:

$$|\Phi(j\omega)|^2 = \left| \frac{k}{j\omega+k} \right|^2 = \frac{k^2}{\omega^2+k^2};$$

$$|\Phi_{em}(j\omega)|^2 = |\Phi_{yf}(j\omega)|^2 = \left| \frac{j\omega}{j\omega+k} \right|^2 = \frac{\omega^2}{\omega^2+k^2}.$$

Выражение для дисперсии ошибки по сигналу $M(t)$ имеет вид

$$D_e^m = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{mm}(\omega) |\Phi_{em}(j\omega)|^2 d\omega.$$

Подынтегральное выражение может быть приведено к виду, удобному для применения табулированных выражений:

$$\begin{aligned} S_{mm}(\omega) |\Phi_{em}(j\omega)|^2 &= \frac{2D\mu}{\omega^2 + \mu^2} \frac{\omega^2}{\omega^2 + k^2} = \\ &= 2D\mu \frac{-(j\omega)^2}{(j\omega + \mu)(-j\omega + \mu)(j\omega + k)(-j\omega + k)} = \\ &= 2D\mu \frac{-(j\omega)^2}{[(j\omega)^2 + (k + \mu)j\omega + k\mu][(-j\omega)^2 + (k + \mu)(-j\omega) + k\mu]} \end{aligned}$$

Тогда, в соответствии с табл. 15.1, получаем

$$D_e^m = 2D\mu J_2 = 2D\mu \frac{-b_1 + b_0 a_2 / a_0}{2a_2 a_1},$$

причем $a_2 = 1$, $a_1 = k + \mu$, $a_0 = k\mu$; $b_1 = -1$, $b_0 = 0$.

Окончательное выражение для дисперсии ошибки системы по полезному сигналу

$$D_e^m = \frac{D\mu}{k + \mu}. \quad (15.6)$$

Вследствие равенства передаточных функций и спектральных плотностей, дисперсия выходного сигнала по возмущению

$$D_y^f = D_e^m.$$

Анализируя полученные результаты, можно отметить следующее. Дисперсия сигналов для любой переменной системы пропорциональна дисперсии входных сигналов D , что объясняется линейными свойствами системы.

При $k \rightarrow 0$, когда контур системы фактически разомкнут, а сигнал ошибки $E(t)$ совпадает с $M(t)$ (при $N(t) = 0$), сигнал на выходе $Y(t)$ совпадает с $F(t)$, получаем

$$D_e^m = D_y^f = D.$$

При $k \rightarrow \infty$ полезный сигнал $M(t)$ воспроизводится на выходе без искажений (при $N(t) = 0$), а сигнал возмущения $F(t)$ полностью подавляется системой. Вследствие этого

$$D_e^m = D_y^f = 0.$$

Увеличение параметра μ ведет к возрастанию доли составляющих в спектрах входных сигналов $M(t)$ и $F(t)$, приходящихся на высокочастотную область. Эти составляющие подавляются системой, что в результате приводит к большему искажению входных сигналов. Следовательно, с ростом μ дисперсии ошибки системы по полезному сигналу $M(t)$ и сигнала на выходе по возмущению $F(t)$ также возрастают и в пределе стремятся к дисперсиям входных сигналов.

При рассмотрении прохождения помехи $N(t)$, представленной идеальным белым шумом, выражение для дисперсии сигнала на выходе системы имеет вид

$$D_y^n = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{nn}(\omega) |\Phi(j\omega)|^2 d\omega.$$

Подынтегральная функция приводится к выражению

$$S_{nn}(\omega) |\Phi(j\omega)|^2 = c^2 \frac{k^2}{\omega^2 + k^2} = c^2 k^2 \frac{1}{(j\omega + k)(-j\omega + k)}.$$

Тогда, используя соответствующую формулу из табл. 15.1, получаем

$$D_y^n = c^2 k^2 J_1 = c^2 k^2 \frac{b_0}{2a_0 a_1},$$

причем $a_1 = 1$, $a_0 = k$, $b_0 = 1$.

Окончательно дисперсия помехи на выходе записывается как

$$D_y^n = c^2 k / 2. \quad (15.7)$$

Следовательно, чем больше значение k , т. е. чем больше частота среза $\omega_{\text{ср}}$ системы, тем более интенсивно помеха проходит на выход, что и приводит в конечном итоге к увеличению дисперсии D_y^n .

15.6.3. Вычисление дисперсии ошибки в системах второго порядка

Будем рассматривать передаточную функцию разомкнутой системы вида

$$W(s) = \frac{k}{s(Ts + 1)}, \quad k > 0, \quad T > 0.$$

Передаточная функция замкнутой системы

$$\Phi(s) = \frac{k}{Ts^2 + s + k}.$$

В отсутствие помехи передаточные функции замкнутой системы по ошибке и по возмущению:

$$\Phi_{em}(s) = \Phi_y(s) = \frac{Ts^2 + s}{Ts^2 + s + k}.$$

Представим квадраты модулей комплексных коэффициентов передач в следующем виде:

$$|\Phi(j\omega)|^2 = \left| \frac{k}{T(j\omega)^2 + j\omega + k} \right|^2 = \frac{k^2}{[T(j\omega)^2 + j\omega + k][T(-j\omega)^2 + (-j\omega) + k]};$$

$$|\Phi_{em}(j\omega)|^2 = \left| \frac{T(j\omega)^2 + j\omega}{T(j\omega)^2 + j\omega + k} \right|^2 = \frac{T^2 \omega^4 + \omega^2}{[T(j\omega)^2 + j\omega + k][T(-j\omega)^2 + (-j\omega) + k]}.$$

Произведение спектральной плотности сигнала $M(t)$ на квадрат модуля комплексного коэффициента передачи по ошибке приводится к виду, удобному для использования табулированного выражения:

$$\begin{aligned}
& S_{mm}(\omega) |\Phi_{em}(j\omega)|^2 = \\
& = \frac{2D\mu}{(j\omega + \mu)(-j\omega + \mu) [T(j\omega)^2 + j\omega + k][T(-j\omega)^2 + (-j\omega) + k]} = \\
& = 2D\mu \frac{T^2(j\omega)^4 - (j\omega)^2}{[T(j\omega)^3 + (1 + \mu T)(j\omega)^2 + (k + \mu)j\omega + k\mu]} \times \\
& \times \frac{1}{[T(-j\omega)^3 + (1 + \mu T)(-j\omega)^2 + (k + \mu)(-j\omega) + k\mu]}.
\end{aligned}$$

Дисперсия ошибки по полезному сигналу

$$D_e^m = 2D\mu J_3 = 2D\mu \frac{-a_1 b_2 + a_3 b_1 - a_3 a_2 b_0 / a_0}{2a_3(a_3 a_0 - a_2 a_1)},$$

причем $b_2 = T^2$, $b_1 = -1$, $b_0 = 0$; $a_3 = T$, $a_2 = 1 + \mu T$, $a_1 = k + \mu$, $a_0 = k\mu$.

После подстановки коэффициентов и соответствующих преобразований в окончательном виде получаем

$$D_e^m = D \frac{k\mu T + \mu + \mu^2 T}{k + \mu + \mu^2 T} = D_y^f. \quad (15.8)$$

Анализ выражения D_e^m для предельных значений параметров позволяет сделать ряд выводов.

При $T \rightarrow 0$ передаточная функция $W(s)$ стремится к передаточной функции интегрирующего звена и выражение для дисперсии D_e^m приближается к ранее полученному для системы первого порядка.

При $\mu T = 1$ дисперсия сигнала ошибки принимает вид $D_e^m = D$, т. е. дисперсия не зависит от коэффициента передачи k и совпадает с дисперсией воздействия.

Представим выражение для дисперсии ошибки как

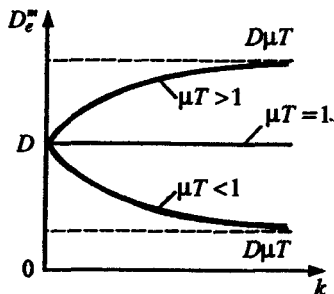
$$D_e^m = D \left[1 + \frac{k(\mu T - 1)}{k + \mu + \mu^2 T} \right].$$

При $\mu T > 1$ числитель D_e^m превышает знаменатель при любом k . С увеличением k происходит рост значения D_e^m и в пределе $D_e^m = D\mu T > D$.

При $\mu T < 1$ числитель D_e^m меньше знаменателя при любом k и $D_e^m < D$. Чем меньше значение μT , тем лучше в системе подавляется возмущение $F(t)$ и воспроизводится полезное задающее воздействие $M(t)$. С увеличением k значение D_e^m уменьшается и в пределе достигает $D\mu T$.

Изменение D_e^m от k в зависимости от μT приведено на рис. 15.11.

Рис. 15.11. Графики зависимости дисперсии ошибки системы по полезному сигналу от коэффициента передачи



Дисперсию сигнала ошибки можно представить так:

$$D_e^m = \frac{D\mu}{k + \mu + \mu^2 T} + \frac{D\mu(k + \mu)T}{k + \mu + \mu^2 T}.$$

Если предположить, что $\mu^2 T \ll k + \mu$ или $\mu T \ll 1 + k/\mu$, то

$$D_e^m \approx \frac{D\mu}{k + \mu} + D\mu T = \frac{D}{1 + (k/\mu)} + D\mu T.$$

Первое слагаемое совпадает с ранее полученным выражением для дисперсии сигнала ошибки в системе первого порядка, отражая фильтрующие свойства системы как низкочастотного фильтра. Второе слагаемое обусловлено влиянием резонансных (колебательных) свойств замкнутой системы, которые вызываются наличием в выражении для $W(s)$ постоянной времени T . В случае, когда $\mu T = 1$, эффект влияния фильтрующих свойств компенсируется влиянием резонансных свойств, а когда $\mu T > 1$ влияние резонансных свойств становится преобладающим [69].

При рассмотрении прохождения помехи $N(t)$ произведение спектральной плотности помехи на квадрат модуля комплексного коэффициента передачи системы приводится к выражению

$$S_{nn}(\omega) |\Phi(j\omega)|^2 = c^2 \left| \frac{k}{T(j\omega)^2 + j\omega + k} \right|^2 = c^2 \frac{k^2}{[T(j\omega)^2 + j\omega + k][T(-j\omega)^2 + (-j\omega) + k]}.$$

Тогда дисперсия помехи на выходе:

$$D_y^n = c^2 k^2 J_2 = c^2 k^2 \frac{-b_1 + b_0 a_2 / a_0}{2a_2 a_1},$$

причем $b_1 = 0$, $b_0 = 1$; $a_2 = T$, $a_1 = 1$, $a_0 = k$. При подстановке значений коэффициентов в D_y^n получаем

$$D_y^* = c^2 k / 2.$$

Полученное выражение для дисперсии помехи на выходе полностью совпадает с выражением (15.7) для системы первого порядка и не зависит от постоянной времени T .

На рис. 15.12 изображены кривые квадрата модуля частотных характеристик замкнутых систем первого (кривая 1) и второго (кривая 2) порядков. В области высоких частот кривая 2 проходит ниже кривой 1, т. е. проявление фильтрующих свойств в системе второго порядка сказывается больше. Однако в области низких частот кривая 2 проходит выше, что связано с проявлением резонансных свойств системы. Чем больше значение T , тем сильнее проявляются фильтрующие свойства системы в области высоких частот и резонансные свойства в области низких частот [69].

Величины заштрихованных на рис. 15.12 площадей S_1 и S_2 одинаковы независимо от значения T , поэтому при воздействии на систему идеального белого шума, имеющего постоянное значение спектральной плотности, в выражении для D_y^* отсутствует постоянная времени T .

Следовательно, при исследовании замкнутых систем управления необходимо учитывать не только их фильтрующие свойства, подавляющие высокочастотные сигналы, но и резонансные, колебательные свойства, проявляющиеся в области низких частот.

При одинаковых значениях D_y^* , определяющих одинаковый разброс значений сигналов на выходах рассматриваемых замкнутых систем первого и второго порядков, спектральный состав сигналов различен. Для системы второго порядка более подчеркнуты низко- и среднечастотные составляющие сигнала по сравнению с первым порядком и в меньшей степени проявляются высокочастотные составляющие.

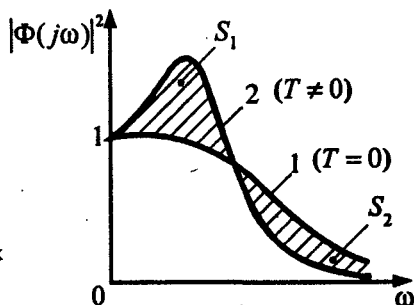


Рис. 15.12. Графики квадрата модулей частотных характеристик для замкнутых систем первого и второго порядков

В том случае, когда помеха $N(t)$ может быть представлена реальным белым шумом (14.5), произведение спектральной плотности помехи на квадрат модуля частотной характеристики замкнутой системы имеет вид:

$$\begin{aligned} S_{nn}(\omega) |\Phi(j\omega)|^2 &= \frac{2D\mu}{\omega^2 + \mu^2} \left| \frac{k}{T(j\omega)^2 + j\omega + k} \right|^2 = \\ &= 2D\mu k^2 \frac{1}{(j\omega + \mu)(-j\omega + \mu)[T(j\omega)^2 + j\omega + k][T(-j\omega)^2 + (-j\omega) + k]} = \\ &= 2D\mu k^2 \frac{1}{[T(j\omega)^3 + (1 + \mu T)(j\omega)^2 + (k + \mu)(j\omega) + k\mu]} \times \\ &\quad \times \frac{1}{[T(-j\omega)^3 + (1 + \mu T)(-j\omega)^2 + (k + \mu)(-j\omega) + k\mu]} \end{aligned}$$

Тогда дисперсия сигнала на выходе, вызванная действием помехи $N(t)$

$$D_y^n = 2D\mu k^2 J_3 = 2D\mu k^2 \frac{-a_1 b_2 + a_3 b_1 - a_3 a_2 b_0 / a_0}{2a_3(a_3 a_0 - a_2 a_1)},$$

причем $b_2 = b_1 = 0$, $b_0 = 1$; $a_3 = T$, $a_2 = 1 + \mu T$, $a_1 = k + \mu$, $a_0 = k\mu$. При подстановке коэффициентов в выражение D_y^n получаем

$$D_y^n = D \frac{k(1 + \mu T)}{k + \mu + \mu^2 T}.$$

В том случае, если $\mu T > kT - 1$, то дисперсия помехи на выходе $D_y^n < D$, т. е. система подавляет помеху $N(t)$.

Следует отметить, что при действии на входе системы (см. рис. 15.10) только одной помехи $N(t)$ со спектральной плотностью (14.5), сумма сигналов на выходе системы $Y(t)$ и на выходе элемента сравнения $E(t)$ определяется выражением

$$E(t) + Y(t) = N(t),$$

их суммарная дисперсия

$$D_\Sigma = D_e^n + D_y^n = D\mu \frac{(k + \mu)T + 1}{k + \mu + \mu^2 T} + D \frac{k(1 + \mu T)}{k + \mu + \mu^2 T} = D \left(1 + \frac{2k\mu T}{k + \mu + \mu^2 T} \right) > D.$$

Превышение суммарной дисперсии D_Σ над дисперсией входного сигнала $N(t) = E(t) + Y(t)$ обусловлено взаимной вероятностной связью (корреляцией) сигналов $E(t)$ и $Y(t)$.

15.6.4. Вычисление дисперсии ошибки в системах третьего порядка

Рассматривается передаточная функция разомкнутой системы:

$$W(s) = \frac{k}{s(T_1 s + 1)(T_2 s + 1)}, \quad k > 0, T_1 > 0, T_2 > 0$$

Передаточная функция замкнутой системы

$$\Phi(s) = \frac{k}{T_1 T_2 s^3 + (T_1 + T_2)s^2 + s + k},$$

передаточные функции замкнутой системы по ошибке $E(t)$ и по возмущению $F(t)$ при $N(t) = 0$:

$$\Phi_{em}(s) = \Phi_{yf}(s) = \frac{T_1 T_2 s^3 + (T_1 + T_2)s^2 + s}{T_1 T_2 s^3 + (T_1 + T_2)s^2 + s + k}.$$

Дисперсия ошибки по полезному сигналу

$$D_e^m = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{mm}(\omega) |\Phi_{em}(j\omega)|^2 d\omega =$$

$$= \frac{D\mu}{\pi} \int_{-\infty}^{\infty} \frac{-(T_1 T_2)^2 (j\omega)^6 + (T_1^2 + T_2^2)(j\omega)^4 - (j\omega)^2}{A(j\omega)A(-j\omega)} d\omega,$$

где $A(j\omega) = T_1 T_2 (j\omega)^4 + (T_1 + T_2 + \mu T_1 T_2)(j\omega)^3 + [1 + \mu(T_1 + T_2)](j\omega)^2 + (\mu + k)(j\omega) + \mu k$.

Используя соответствующую формулу из табл. 15.1, получаем

$$D_e^m = 2D\mu J_4 = 2D\mu \frac{b_3(-a_3 a_0 + a_2 a_1) - a_4 a_1 b_2 + a_4 a_3 b_1 + a_4 b_0(a_4 a_1 - a_3 a_2)/a_0}{2a_4(a_4 a_1^2 + a_3^2 a_0 - a_3 a_2 a_1)},$$

причем

$$b_3 = (T_1 T_2)^2, \quad b_2 = T_1^2 + T_2^2, \quad b_1 = -1, \quad b_0 = 0; \quad a_4 = T_1 T_2, \quad a_3 = T_1 + T_2 + \mu T_1 T_2,$$

$$a_2 = 1 + \mu(T_1 + T_2), \quad a_1 = \mu + k, \quad a_0 = \mu k.$$

Подставляя коэффициенты в выражение для дисперсии, можно записать

$$D_e^m = D\mu \frac{(T_1 T_2 \mu^2 + 1)(T_1 + T_2 - T_1 T_2 k) + (\mu + k)(T_1 + T_2)^2}{(T_1 + T_2 - T_1 T_2 k)[k + \mu + \mu^2(T_1 + T_2 + \mu T_1 T_2)]} = D_y^f.$$

При $k = 0$ получаем $D_e^m = D_y^f = D$, что согласуется с физическим толкованием результата.

При $T_1 = 0$ или $T_2 = 0$ выражение для дисперсии сигнала ошибки совпадает с ранее найденным (15.8) для системы второго порядка.

При $k = (T_1 + T_2)/T_1 T_2 = k_{\text{кр}}$, когда система находится на границе устойчивости, а модуль частотной характеристики замкнутой системы на частоте $\omega = 1/\sqrt{T_1 T_2}$ достигает бесконечности, дисперсия D_e^m также принимает бесконечное значение.

В том случае, когда

$$k < k_{\text{кр}} [1 - \mu(T_1 + T_2) - \mu^2 T_1 T_2],$$

дисперсия сигнала ошибки отвечает неравенству

$$D_e^m < D.$$

При этом, чтобы $k > 0$, должно выполняться условие

$$\mu(T_1 + T_2 + \mu T_1 T_2) < 1. \quad (15.9)$$

Только при выполнении условия (15.9) в определенном интервале изменения коэффициента передачи k дисперсия ошибки системы по полезному сигналу может принимать значения меньше дисперсии входного сигнала, а следовательно, в системе обеспечивается воспроизведение на выходе полезного сигнала $M(t)$ и подавление возмущения $F(t)$.

Представим дисперсию D_e^m в следующем виде

$$D_e^m = D\mu \left\{ \frac{1 + \mu^2 T_1 T_2}{k + \mu + \mu^2 (T_1 + T_2 + \mu T_1 T_2)} + \frac{(\mu + k)(T_1 + T_2)^2}{(T_1 + T_2 - k T_1 T_2)[k + \mu + \mu^2 (T_1 + T_2 + \mu T_1 T_2)]} \right\}.$$

Тогда, при выполнении «сильных» неравенств:

$$\mu^2 T_1 T_2 \ll 1; \quad \mu^2 (T_1 + T_2 + \mu T_1 T_2) \ll k + \mu$$

можно приближенно записать

$$D_e^m \approx \frac{D\mu}{k + \mu} + \frac{D\mu(T_1 + T_2)}{1 - k T_1 T_2 / (T_1 + T_2)}.$$

Первое слагаемое отражает фильтрующие, второе — резонансные свойства замкнутой системы.

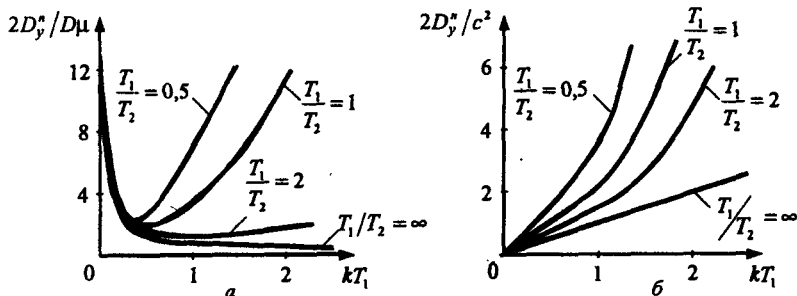


Рис. 15.13. Графики зависимости дисперсии ошибки по полезному сигналу (а) и дисперсии помехи на выходе (б) от коэффициента передачи и постоянных времени

Дисперсия выходного сигнала вследствие действия только помехи $N(t)$ в виде идеального белого шума:

$$\begin{aligned}
 D_y^n &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{nn}(\omega) |\Phi(j\omega)|^2 d\omega = \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} c^2 \left| \frac{k}{T_1 T_2 (j\omega)^3 + (T_1 + T_2)(j\omega)^2 + j\omega + k} \right|^2 d\omega = \\
 &= k^2 c^2 J_3 = k^2 c^2 \frac{-a_1 b_2 + a_3 b_1 - a_3 a_2 b_0 / a_0}{2a_3 (a_3 a_0 - a_2 a_1)},
 \end{aligned}$$

причем $b_2 = b_1 = 0$, $b_0 = 1$; $a_3 = T_1 T_2$, $a_2 = T_1 + T_2$, $a_1 = 1$, $a_0 = k$. При подстановке значений коэффициентов в выражение для дисперсии, получаем

$$D_y^n = \frac{kc^2}{2} \frac{1}{1 - kT_1 T_2 / (T_1 + T_2)}.$$

Семейство графиков $D_y^n/D\mu$ и $2D_y^n/c^2$ в зависимости от коэффициента передачи k и постоянных времени T_1 и T_2 при заданном $\mu = 0,1$ приведены на рис. 15.13. В отличие от рассмотренных ранее систем первого и второго порядков, зависимость D_y^n от k при фиксированных значениях T_1 и T_2 носит нелинейный характер. С увеличением k наблюдается рост D_y^n , что можно объяснить увеличением резонансных свойств системы по сравнению с фильтрующими свойствами [69].

СИНТЕЗ ЛИНЕЙНЫХ СИСТЕМ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

16.1. ПОСТАНОВКА ЗАДАЧ СИНТЕЗА

Будем рассматривать систему управления (рис. 16.1) с передаточной функцией $W(s)$ как преобразователь входной смеси $G(t)$ полезного сигнала $M(t)$ и помехи $N(t)$ в выходной случайный сигнал $Y(t)$. Идеальное преобразование полезного сигнала $M(t)$ в желаемый сигнал $Z(t)$ осуществляется заданным оператором (алгоритмом преобразования) $H(s)$.

В самом общем случае требуется возможно точнее воспроизвести на выходе системы сигнал $Z(t)$ или, другими словами, свести к минимуму ошибку $\Theta(t) = Z(t) - Y(t)$ между желаемым $Z(t)$ и реальным $Y(t)$ сигналами.

Задачу отделения полезного сигнала (случайного или регулярного) от помехи и подавления (фильтрации) последней называют *задачей фильтрации* или *сглаживания*.

Введение оператора преобразования $H(s)$ позволяет распространить задачу фильтрации не только на следящие (воспроизводящие) системы, для которых $Z(t) \equiv M(t)$ (т. е. $H(s) \equiv 1$), но и на другие виды систем. В зависимости от вида оператора $H(s)$ возможна классификация линейных систем, которая представлена в табл. 16.1.

Идеальное преобразование полезного сигнала невозможно из-за динамических свойств системы, а также вследствие возмущающих воздействий (помех).

Синтез систем при случайных воздействиях состоит в определении структур и параметров операторов преобразования сигналов (динамических характеристик), обеспечивающих наилучшим образом выполнение заданного вероятностного критерия качества (оптимальности).

Качество функционирования систем управления оценивается по различным критериям. Одним из наиболее важных является критерий, характеризующий точность преобразования, обычно воспроизведения, по-

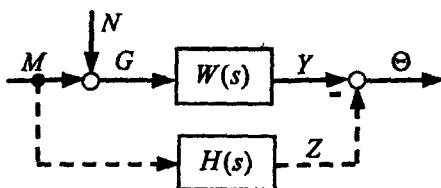


Рис. 16.1. Реальное и идеальное преобразования полезного сигнала

Название системы	Алгоритм преобразования	Оператор преобразования $H(s)$
Масштабирующая	$Z(t) = kM(t)$	$k = \text{const}$
Упреждающая (предсказывающая, экстраполирующая)	$X(t) = M(t + \tau)$	$e^{s\tau}, \tau > 0$
Запаздывающая	$Z(t) = M(t - \tau)$	$e^{-s\tau}$
Интегрирующая	$Z(t) = \int_0^t M(t) dt$	$1/s$
Дифференцирующая	$Z(t) = dM(t)/dt$	s

лезных сигналов в условиях действия различных случайных возмущений. На практике часто за вероятностный критерий оптимальности принимают критерий минимума среднеквадратической ошибки $\theta_{\text{ср}}$ или, при отсутствии регулярной составляющей, критерий минимума дисперсии ошибки $D_\theta = \theta_{\text{ср}}^2$.

В этом случае задача синтеза сводится к нахождению такой физически реализуемой оптимальной передаточной функции $W_{\text{опт}}(s)$ системы, при которой $D_\theta = \min$.

При использовании критериев минимума дисперсии ошибки D_θ или среднеквадратической ошибки $\theta_{\text{ср}}$ оценка точности системы производится в зависимости от среднего, а не от мгновенного значения ошибки. Критерий не применим, когда требуется, чтобы ошибка не выходила за заданные ограничения, либо, когда требования к величине ошибки в разные моменты времени неодинаковы.

Если бы к системе было приложено только одно внешнее воздействие в виде полезного сигнала $G(t) = M(t)$, либо помехи $G(t) = N(t)$, то соответствующим выбором параметров передаточных функций звеньев системы теоретически можно было бы обеспечить любую точность системы. Однако при одновременном действии полезного сигнала и помехи достигаемая точность системы не может быть любой.

На рис. 16.2 приведены типичные графики зависимости составляющих дисперсии ошибки от коэффициента передачи k разомкнутой систе-

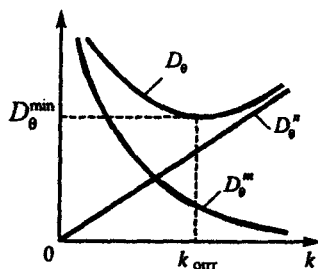


Рис. 16.2. Графики зависимости составляющих дисперсии ошибки от коэффициента передачи системы

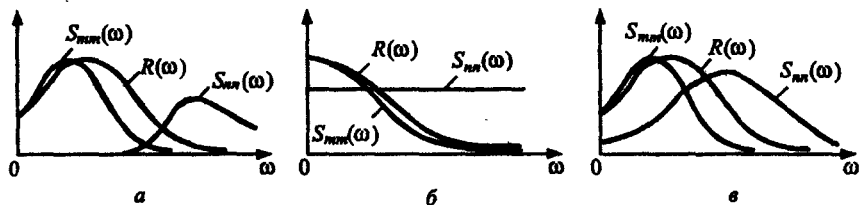


Рис. 16.3. Взаимное расположение графиков спектральных плотностей воздействий и амплитудно-частотной характеристики системы

мы. Как следует из анализа кривых, для лучшего воспроизведения полезного сигнала $M(t)$, соответственно, уменьшения D_{θ}^* , система должна иметь возможно больший коэффициент передачи. Для лучшего подавления помехи $N(t)$, т. е. уменьшения D_{θ}^* система должна иметь возможно меньший коэффициент k . При одновременном действии двух воздействий возможно получение оптимального решения $k = k_{\text{опт}}$, при котором $D_{\theta}^{\min} < D_{\theta}^*$.

В том случае, когда спектральные плотности полезного сигнала $S_{\text{сиг}}(\omega)$ и помехи $S_{\text{пом}}(\omega)$ достаточно разнесены друг от друга (рис. 16.3, а), амплитудно-частотную характеристику $R(\omega) = |W(j\omega)|$ системы выбирают достаточно широкой для пропускания спектральных составляющих сигнала $M(t)$ (с целью обеспечения требуемой точности воспроизведения полезного воздействия) и достаточно узкой для подавления спектральных составляющих сигнала $N(t)$ (с целью уменьшения реагирования системы на помеху).

На рис. 16.3, б изображены графики спектральных плотностей полезного сигнала и помехи типа белого шума. Основная часть спектральной плотности (мощности) полезного сигнала сконцентрирована в области низких частот, поэтому низкочастотная часть амплитудно-частотной характеристики $R(\omega)$ должна выбираться возможно близкой к форме спектральной плотности $S_{\text{сиг}}(\omega)$, а ее высокочастотная часть должна быстро убывать, по возможности следуя за убывающей характеристикой $S_{\text{пом}}(\omega)$.

В общем случае, когда кривые спектральных плотностей сигналов накладываются друг на друга и имеют произвольную форму (рис. 16.3, в), определение оптимальных параметров системы в значительной степени усложняется. При этом синтез систем управления в условиях приложения случайных воздействий допускает две постановки задачи.

1. *Синтез оптимальной системы при заданной структуре и известных вероятностных характеристиках воздействий*, когда в результате выбора параметров системы обеспечивается минимум дисперсии ошибки.

2. Синтез оптимальной системы при произвольной структуре и известных вероятностных характеристиках воздействий, когда в результате выбора структуры и параметров системы обеспечивается минимум дисперсии ошибки.

16.2. СИНТЕЗ ОПТИМАЛЬНОЙ СИСТЕМЫ С ЗАДАННОЙ СТРУКТУРОЙ

Будем считать, что структура системы (см. рис. 16.1), определяемая взаимосвязью звеньев и видом оператора $W(s)$, а также вероятностные характеристики внешних воздействий полностью известны. Часть параметров $q_1, \dots, q_r$ звеньев системы допускают варьирование в определенных пределах. Тогда, за счет изменения варьируемых параметров возможно найти минимальное значение дисперсии ошибки, используя в качестве критерия оптимальности выражение (15.5).

Нахождение вектора параметров $\mathbf{q} = (q_1, \dots, q_r)$, при которых

$$\min_{\mathbf{q}} D_{\theta}(\mathbf{q}) = D_{\theta}^{\min}$$

и составляет содержание задачи синтеза системы с заданной структурой, которая по сути представляет собой задачу параметрической оптимизации (см. § 4.9).

При решении поставленной задачи предполагается, что минимум функции $D_{\theta}(\mathbf{q})$ в заданных пределах изменения варьируемых параметров существует и единственен. Тогда вектор оптимальных значений параметров $\mathbf{q}$ определяется в результате решения уравнения

$$\nabla D_{\theta}(\mathbf{q}) = 0, \quad (16.1)$$

где $\nabla D_{\theta}(\mathbf{q}) = \text{grad} D_{\theta}(\mathbf{q}) = \left(\frac{\partial D_{\theta}(\mathbf{q})}{\partial q_1}, \dots, \frac{\partial D_{\theta}(\mathbf{q})}{\partial q_r} \right)$ — вектор-градиент функции $D_{\theta}(\mathbf{q})$.

Во многих случаях, когда порядок системы высокий, выражения для частных производных имеют достаточно сложный вид. Определение минимума D_{θ} производится с использованием различных численных процедур поиска экстремума функции.

При решении задачи оптимизации следует в первую очередь установить возможность получения решения при заданной структуре системы и приложенных воздействиях.

Например, пусть все передаточные функции системы, представленной на рис. 16.4, физически осуществимы, причем степень полинома числителя $W_2(s)$ меньше степени полинома знаменателя. В этом случае,

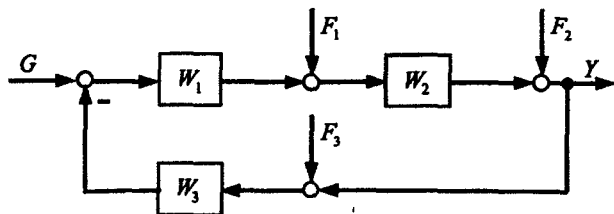


Рис. 16.4. Выбор моделей воздействий в зависимости от мест их приложения в системе

воздействие $F_2(t)$ не может быть представлено идеальным белым шумом, так как комплексный коэффициент передачи $\Phi_{y_2}(j\omega)$ соответствует высокочастотному фильтру. В то же время воздействия $G(t)$, $F_1(t)$, $F_3(t)$ допускают представление в виде идеального белого шума.

Для выполнения оптимизационной процедуры на основе решения уравнения (16.1), как отмечалось в § 16.1, необходимо, чтобы при возрастании варьируемого параметра одни составляющие дисперсии ошибки увеличивались, а другие, напротив, уменьшались, либо хотя бы одна из составляющих дисперсии обладала бы минимумом по этому параметру.

В частности, если для системы с передаточной функцией $W(s) = k/s$, структура которой приведена к расчетному виду (см. рис. 15.10), составляющие дисперсии ошибки D_e^m и D_e^y монотонно убывают, а D_e^y — возрастает с увеличением коэффициента передачи k , то оптимизационная постановка задачи возможна лишь при учете по крайней мере двух воздействий $M(t)$ и $N(t)$, либо $F(t)$ и $N(t)$.

Учет же только двух воздействий $M(t)$ и $F(t)$ исключает возможность получения оптимального решения (16.1), так как обе составляющие дисперсии ошибки, вызываемые действием этих сигналов, одновременно убывают с ростом k . Очевидно оптимизационная постановка задачи невозможна при учете лишь одного воздействия.

Следует отметить, что и при соблюдении вышеприведенных условий зависимость дисперсии ошибки $D_e(k)$ системы может не иметь выраженного минимума. Например, в случае, когда дисперсия ошибки относительно полезного воздействия M и помехи N_1 с интенсивностью c_1^2 имеет вид

$$D_{e_1} = D_{e_1}^m + D_{e_1}^{n_1} = \frac{D\mu}{k + \mu} + \frac{kc_1^2}{2},$$

причем $D/\mu < c_1^2/2$, зависимость $D_{e_1}(k)$ представляет собой монотонно возрастающую функцию, минимальное значение которой достигается при $k = 0$. Это объясняется высоким уровнем интенсивности помехи по отношению к интенсивности полезного воздействия. Увеличение зна-

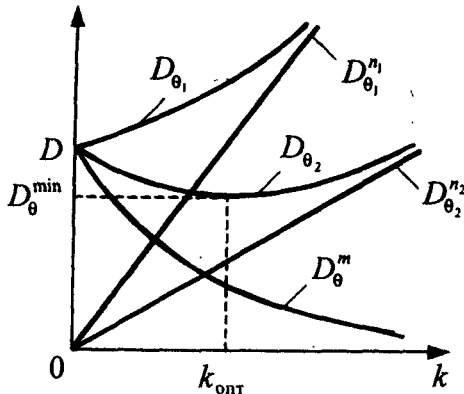


Рис. 16.5. Графики дисперсии ошибки системы при различных интенсивностях помехи

чения k лишь в большей степени подчеркивает помеху на фоне слабого полезного воздействия.

В случае, когда интенсивность c_2^2 помехи N_2 на фоне полезного воздействия такова, что $D/\mu > c_2^2/2$, зависимость $D_{\theta_2}(k)$ имеет выраженный минимум, отвечающий оптимальному значению коэффициента передачи.

Зависимости $D_{\theta_1}(k)$, $D_{\theta_1}^{n1}(k)$ и $D_{\theta_2}(k)$, $D_{\theta_2}^{n2}(k)$, обусловленные помехами N_1 и N_2 соответственно, приведены на рис. 16.5.

При оптимизации систем первого или второго порядков, когда к системе прилагается лишь одно случайное воздействие, минимальное значение дисперсии ошибки достигается либо при нулевом, либо при бесконечном значении k .

Для систем, порядок которых выше двух, минимум дисперсии ошибки может достигаться и при одном случайном воздействии [69]. В этом случае при малых значениях коэффициента передачи k системы дисперсия ошибки близка к дисперсии входного сигнала. С увеличением значения коэффициента передачи при определенных соотношениях между диапазоном частот, в котором сосредоточен спектр полезного воздействия, и постоянными времени системы величина дисперсии ошибки понижается, а начиная со значений коэффициента, когда повышается колебательность системы, дисперсия ошибки возрастает и достигает бесконечного значения при критическом значении коэффициента передачи.

Решение оптимизационной задачи возможно и при одновременном действии регулярного полезного сигнала, регулярного возмущения и случайной помехи, когда при изменении варьируемого параметра системы, величина составляющей ошибки, вызываемой регулярными воз-

действиями, и величина составляющей ошибки, вызываемой случай- ной помехой, изменяются в разных направлениях.

В случае приложения к системе регулярных воздействий устано- вившаяся ошибка является регулярной функцией. Для степенных воз- действий установившаяся ошибка зависит от степени воздействия и порядка астатизма системы (см. п. 3.6.2).

Пример 1. Пусть рассматривается поведение системы (см. рис. 15.10) с переда- точной функцией $W(s) = k/s$, на вход которой поступает аддитивная случайная смесь сигналов $M(t)$ и $N(t)$, причем $M(t)$ — полезный сигнал со спектральной плотностью $S_{mm}(\omega) = 2D\mu/(\omega^2 + \mu^2)$, $N(t)$ — помеха со спектральной плотностью $S_{nn}(\omega) = c^2$. Воз- мущение $F(t)$ отсутствует.

Дисперсия ошибки

$$D_{\theta} = D_e^m + D_y^n.$$

Требуется определить коэффициент передачи $k = k_{\text{опт}}$, при котором обеспечива- ется минимум D_{θ} .

Согласно (15.6), (15.7) можно записать выражение для дисперсии

$$D_{\theta} = \frac{D\mu}{k + \mu} + c^2 \frac{k}{2}.$$

Частная производная

$$\frac{\partial D_{\theta}}{\partial k} = -\frac{D\mu}{(k + \mu)^2} + \frac{c^2}{2}.$$

Приравнявая к нулю частную производную, получаем уравнение

$$c^2 k^2 + 2\mu c^2 k + \mu^2 c^2 - 2D\mu = 0.$$

Решение уравнения относительно коэффициента передачи:

$$k_{1,2} = -\mu \pm \sqrt{2D\mu}/c.$$

Для системы с отрицательной обратной связью значение k должно быть поло- жительным, следовательно

$$k_{\text{опт}} = \sqrt{2D\mu}/c - \mu.$$

Условие $k_{\text{опт}} > 0$ требует выполнения неравенства

$$\frac{D}{\mu} > \frac{c^2}{2}.$$

Минимальное значение дисперсии

$$D_{\theta}^{\min} = c(\sqrt{2D\mu} - \mu c/2). \quad (16.2)$$

Вводя обозначение $(\mu c^2/2)/D = \alpha$, в соответствии с полученным неравенством, имеем $\alpha < 1$. Тогда

$$D_{\theta}^{\min} = (2\sqrt{\alpha} - \alpha)D.$$

$$\delta_{D_0} = D_{\theta}^{\min} / D = \sqrt{\alpha} (2 - \sqrt{\alpha})$$

характеризует эффективность фильтрации. Величина δ_{D_0} равна нулю при $\alpha = 0$ и возрастает с увеличением α вплоть до значения $\alpha = 1$.

Если требуется обеспечить δ_{D_0} , не превосходящей допустимого значения δ , т. е.

$$\delta_{D_0} = \sqrt{\alpha} (2 - \sqrt{\alpha}) < \delta,$$

то необходимо, чтобы для α выполнялось условие

$$\alpha < 2 - \delta - 2\sqrt{1 - \delta}.$$

Например, если $\delta = 0,75$, то $\alpha < 0,25$. При этом $c^2 < 0,5$ (D/μ).

Получаемые в процессе синтеза оптимальные параметры системы необходимо согласовывать с возможностью их технической реализации, а также с другими показателями качества: временем регулирования, перерегулированием и т. д.

16.3. СИНТЕЗ ОПТИМАЛЬНОЙ СИСТЕМЫ С ПРОИЗВОЛЬНОЙ СТРУКТУРОЙ

16.3.1. Описание оптимальной системы интегральным уравнением Винера-Хопфа

Перейдем к рассмотрению системы управления произвольной структуры, на вход которой подается некоррелированная смесь центрированного случайного полезного сигнала $M(t)$ и центрированной помехи $N(t)$ с известными вероятностными характеристиками (см. рис. 16.1).

Система считается оптимальной, если для заданных сигналов $M(t)$ и $N(t)$ ее динамическая характеристика (оператор преобразования) обеспечивает минимум дисперсии ошибки, т. е.

$$D_0 = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \theta^2(t) dt = \min ,$$

где $\theta(t)$ — любая реализация случайной ошибки $\Theta(t)$.

Эта задача впервые, независимо друг от друга, была поставлена и решена Н. Винером и А.Н. Колмогоровым.

В качестве динамической характеристики системы будем использовать либо весовую функцию $w(t)$, либо комплексный коэффициент передачи $W(j\omega)$, однозначно связанный с передаточной функцией $W(s) = L\{w(t)\}$.

Для стационарных эргодических случайных сигналов $M(t)$ и $N(t)$ оптимальная весовая функция $w_{\text{опт}}(t)$ ищется в классе весовых функций линейных стационарных систем, т. е. функций, зависящих от одного аргумента времени.

На весовую функцию $w(t)$ накладывается ограничение в виде условия физической реализуемости (осуществимости) системы:

$$w(t) \equiv 0 \quad \text{при } t < 0.$$

Это условие означает, что реакция на выходе системы не может наступить ранее приложенного к ней входного воздействия.

Сигнал на выходе системы в установившемся режиме равен

$$y(t) = \int_{-\infty}^{\infty} g(t-\lambda)w(\lambda)d\lambda.$$

Тогда установившаяся ошибка системы будет определяться следующим образом:

$$\theta(t) = z(t) - \int_{-\infty}^{\infty} g(t-\lambda)w(\lambda)d\lambda.$$

Дисперсию ошибки можно представить так:

$$\begin{aligned} D_{\theta} &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \left[z(t) - \int_{-\infty}^{\infty} g(t-\lambda)w(\lambda)d\lambda \right]^2 dt = \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T z^2(t)dt - 2 \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T z(t)dt \int_{-\infty}^{\infty} g(t-\lambda)w(\lambda)d\lambda + \\ &+ \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T dt \int_{-\infty}^{\infty} g(t-\lambda)w(\lambda)d\lambda \int_{-\infty}^{\infty} g(t-\eta)w(\eta)d\eta. \end{aligned}$$

Меняя порядок интегрирования и группируя функции $g(t)$ и $z(t)$ так, чтобы можно было записать вероятностные характеристики сигналов, получаем:

$$\begin{aligned} D_{\theta} &= K_{zz}(0) - 2 \int_{-\infty}^{\infty} w(\lambda)d\lambda \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g(t-\lambda)z(t)dt + \\ &+ \int_{-\infty}^{\infty} w(\lambda)d\lambda \int_{-\infty}^{\infty} w(\eta)d\eta \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T g(t-\lambda)g(t-\eta)dt = \\ &= K_{zz}(0) - 2 \int_{-\infty}^{\infty} w(\lambda)K_{zg}(\lambda)d\lambda + \int_{-\infty}^{\infty} w(\lambda)d\lambda \int_{-\infty}^{\infty} w(\eta)K_{gg}(\lambda-\eta)d\eta. \end{aligned}$$

Полученное выражение связывает между собой дисперсию ошибки, вероятностные характеристики входных сигналов и весовую функцию системы, которая должна быть такой, чтобы D_0 была минимальной.

Дисперсия ошибки D_0 представляет собой функционал весовой функции $w(t)$, т. е.

$$D_0 = J \{w(t)\}.$$

Определение оптимальной весовой функции $w(t) = w_{\text{опт}}(t)$ осуществляется на основе решения уравнения

$$\left. \frac{\partial J \{w(\lambda) + \gamma \delta w(\lambda)\}}{\partial \gamma} \right|_{\gamma=0} = 0, \quad (16.3)$$

где $\delta w(\lambda)$ — приращение $w(\lambda)$, отвечающее условию физической реализуемости.

Напомним, что оптимум функции $f(t)$ отыскивается на основе решения уравнения

$$\frac{df(t)}{dt} = 0. \quad (16.4)$$

Существует и другой способ нахождения оптимума функции $f(t)$, который основан на решении уравнения, аналогичного (16.3):

$$\left. \frac{\partial f(t + \alpha \Delta t)}{\partial \alpha} \right|_{\alpha=0} = 0. \quad (16.5)$$

Можно показать, что уравнение (16.5) эквивалентно (16.4). Для этого представим $f(t + \alpha \Delta t)$ в виде ряда Тейлора:

$$f(t + \alpha \Delta t) = f(t) + \alpha \Delta t f'(t) + \frac{(\alpha \Delta t)^2}{2!} f''(t) + \frac{(\alpha \Delta t)^3}{3!} f'''(t) + \dots$$

Дифференцируя обе части равенства по параметру α

$$\frac{\partial f(t + \alpha \Delta t)}{\partial \alpha} = \Delta t f'(t) + \alpha (\Delta t)^2 f''(t) + \frac{\alpha^2 (\Delta t)^3}{2!} f'''(t) + \dots$$

и полагая $\alpha = 0$, получаем

$$\left. \frac{\partial f(t + \alpha \Delta t)}{\partial \alpha} \right|_{\alpha=0} = \Delta t \frac{df(t)}{dt} = 0,$$

что совпадает с (16.4).

В соответствии с (16.3) подставим в выражение для D_0 вместо $w(\lambda)$ функцию $w(\lambda) + \gamma \delta w(\lambda)$, тогда

$$\begin{aligned}
J\{w(\lambda) + \gamma \delta w(\lambda)\} &= K_{zz}(0) - 2 \int_{-\infty}^{\infty} [w(\lambda) + \gamma \delta w(\lambda)] K_{zg}(\lambda) d\lambda + \\
&+ \int_{-\infty}^{\infty} [w(\lambda) + \gamma \delta w(\lambda)] d\lambda \int_{-\infty}^{\infty} [w(\eta) + \gamma \delta w(\eta)] K_{gg}(\lambda - \eta) d\eta = \\
&= J_0 + \gamma J_1 + \gamma^2 J_2,
\end{aligned} \tag{16.6}$$

где $J_0 = K_{zz}(0) - 2 \int_{-\infty}^{\infty} w(\lambda) K_{zg}(\lambda) d\lambda + \int_{-\infty}^{\infty} w(\lambda) d\lambda \int_{-\infty}^{\infty} w(\eta) K_{gg}(\lambda - \eta) d\eta$;

$$\begin{aligned}
J_1 &= -2 \int_{-\infty}^{\infty} \delta w(\lambda) K_{zg}(\lambda) d\lambda + \int_{-\infty}^{\infty} \delta w(\lambda) d\lambda \int_{-\infty}^{\infty} w(\eta) K_{gg}(\lambda - \eta) d\eta + \\
&+ \int_{-\infty}^{\infty} w(\lambda) d\lambda \int_{-\infty}^{\infty} \delta w(\eta) K_{gg}(\lambda - \eta) d\eta; \\
J_2 &= \int_{-\infty}^{\infty} \delta w(\lambda) d\lambda \int_{-\infty}^{\infty} \delta w(\eta) K_{gg}(\lambda - \eta) d\eta.
\end{aligned}$$

Производная выражения (16.6) по параметру γ :

$$\frac{\partial J\{w(\lambda) + \gamma \delta w(\lambda)\}}{\partial \gamma} = J_1 + 2\gamma J_2.$$

В таком случае условие (16.3) сводится к условию $J_1 = 0$ или в развернутом виде

$$\begin{aligned}
-2 \int_{-\infty}^{\infty} \delta w(\lambda) K_{zg}(\lambda) d\lambda + \int_{-\infty}^{\infty} \delta w(\lambda) d\lambda \int_{-\infty}^{\infty} w(\eta) K_{gg}(\lambda - \eta) d\eta + \\
+ \int_{-\infty}^{\infty} w(\lambda) d\lambda \int_{-\infty}^{\infty} \delta w(\eta) K_{gg}(\lambda - \eta) d\eta = 0.
\end{aligned}$$

Изменяя порядок интегрирования в последнем интеграле, получаем

$$2 \int_{-\infty}^{\infty} \delta w(\lambda) \left[\int_{-\infty}^{\infty} w(\eta) K_{gg}(\lambda - \eta) d\eta - K_{zg}(\lambda) \right] d\lambda = 0. \tag{16.7}$$

Учитывая, что $\delta w(\lambda) \equiv 0$ при $\lambda < 0$, для выполнения условия (16.7) достаточно, чтобы

$$\int_{-\infty}^{\infty} w(\eta) K_{gg}(\lambda - \eta) d\eta - K_{zg}(\lambda) = 0, \quad \lambda \geq 0.$$

Заменяя λ на τ , а η на λ , получаем в привычных обозначениях *интегральное уравнение*, решением которого является оптимальная весовая функция системы. Поскольку $w(\lambda) \equiv 0$ для $\lambda < 0$, окончательно имеем

$$K_{zg}(\tau) - \int_0^{\infty} w(\lambda) K_{gg}(\tau - \lambda) d\lambda = 0, \quad \tau \geq 0. \quad (16.8)$$

Это интегральное уравнение получило название *уравнения Винера-Хопфа* (N. Wiener-E. Hopf).

16.3.2. Определение оптимальной передаточной функции с учетом физической реализуемости системы (фильтр Винера)

Систему управления с весовой функцией, найденной на основе уравнения Винера-Хопфа, часто называют *фильтром Винера*. Основная трудность решения этого уравнения состоит в том, что весовая функция $w(t)$ должна удовлетворять условию физической реализуемости. С целью преодоления указанной трудности представим (16.8) в следующем виде:

$$K_{zg}(\tau) - \int_{-\infty}^{\infty} w(\lambda) K_{gg}(\tau - \lambda) d\lambda = A(\tau), \quad (16.9)$$

причем неизвестная функция $A(\tau) \equiv 0$ при $\tau \geq 0$.

Взаимная корреляционная функция $K_{zg}(\tau)$ желаемого $Z(t)$ и входного $G(t)$ сигналов, а также взаимная корреляционная функция реального сигнала $Y(t)$ на выходе системы и входного сигнала $G(t)$

$$K_{yg}(\tau) = \int_{-\infty}^{\infty} w(\lambda) K_{gg}(\tau - \lambda) d\lambda$$

при $|\tau| \rightarrow \infty$ стремятся к нулю. Следовательно, имеет место предельное соотношение

$$\lim_{\tau \rightarrow -\infty} A(\tau) = 0. \quad (16.10)$$

Возможный вид функции $A(\tau)$ приведен на рис. 16.6.

В соответствии с условием физической реализуемости изображение Фурье-функции $A(\tau)$ не должно иметь нули и полюсы в верхней полуплоскости комплексной плоскости. Если, например, $A(\tau)$ можно представить в виде

$$A(\tau) = \sum_{i=1}^n c_i e^{\alpha_i \tau}.$$

то все α_i должны быть больше нуля, т. е. $\alpha_i \geq 0$, иначе не будет выполняться предельное соотношение (16.10).

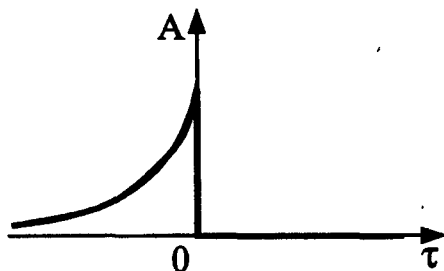


Рис. 16.6. График функции $A(\tau)$

Преобразование Фурье-функции $A(\tau)$ дает

$$F\{A(\tau)\} = \sum_{i=1}^n \frac{c_i}{j\omega_i - \alpha_i},$$

т. е. полюсы изображения этой функции $\omega_i = -j\alpha_i$ лежат в нижней полуплоскости комплексной плоскости ω .

Преобразуем по Фурье уравнение (16.9):

$$S_{zg}(j\omega) - W(j\omega)S_{gg}(\omega) = A(j\omega). \quad (16.11)$$

Потребуем, чтобы комплексный коэффициент передачи $W(j\omega)$, по которому однозначно определяется оптимальная передаточная функция $W_{\text{опт}}(s)$ устойчивой системы, не имел нули и полюсы в нижней полуплоскости. Действительно, любому корню, лежащему в нижней полуплоскости комплексной плоскости, будет соответствовать корень в правой полуплоскости плоскости s . Так, если $\omega_i = -j\alpha_i$, то $s_i = j\omega_i = \alpha_i$. Комплексному коэффициенту передачи $W(j\omega)$ с полюсами в нижней полуплоскости отвечает весовая функция $w(t)$, которая существует и для отрицательных значений t , т. е. не отвечает условию физической реализуемости.

Для определения физически реализуемого комплексного коэффициента передачи $W(j\omega)$ можно воспользоваться методикой, предложенной Г. Бодом и К. Шенноном (С. Shannon). С этой целью введем обозначение $W^*(j\omega)$, означающее, что нули и полюсы $W(j\omega)$ лежат в верхней полуплоскости плоскости ω . Очевидно, изображение $A(j\omega)$, не имеющее нулей и полюсов в верхней полуплоскости, должно быть обозначено как $A^-(j\omega)$.

Представим спектральную плотность входного сигнала в виде произведения:

$$S_{gg}(\omega) = S_{gg}^+(j\omega)S_{gg}^-(j\omega). \quad (16.12)$$

Функция $S_{gg}^+(j\omega)$ не имеет нулей и полюсов в нижней полуплоскости, а $S_{gg}^-(j\omega)$ — в верхней. Такое представление $S_{gg}(j\omega)$ называется *факторизацией*.

С учетом (16.12) уравнение (16.11) можно записать следующим образом:

$$S_{gg}^-(j\omega) \left[\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} - W^+(j\omega) S_{gg}^+(j\omega) \right] = A^-(j\omega). \quad (16.13)$$

Первое слагаемое в квадратных скобках в (16.13) допускает представление в виде

$$\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} = \left[\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} \right]^+ + \left[\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} \right]^-$$

В этом выражении первое слагаемое не имеет нулей и полюсов в нижней полуплоскости, а второе — в верхней. Произведенная операция называется *сепарацией* или *расщеплением*.

Вследствие того, что $A^-(j\omega)$ не имеет нулей и полюсов в верхней полуплоскости, левая часть уравнения (16.13) также не допускает расположения нулей и полюсов в верхней полуплоскости. Это утверждение возможно лишь в том случае, когда

$$\left[\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} \right]^+ - W^+(j\omega) S_{gg}^+(j\omega) = 0.$$

В результате получаем, что оптимальный комплексный коэффициент передачи системы имеет вид

$$W_{\text{опт}}(j\omega) = \left[\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} \right]^+ / S_{gg}^+(j\omega). \quad (16.14)$$

Реализуемая передаточная функция, соответствующая (16.14), обеспечивает в системе минимальное значение дисперсии ошибки для заданных входных стационарных случайных сигналов в установившемся режиме.

Пример 2. Пусть на вход системы поступает аддитивная смесь сигналов $G(t) = M(t) + N(t)$, причем спектральные плотности полезного случайного сигнала $M(t)$ и помехи $N(t)$ равны:

$$S_{mm}(\omega) = \frac{2D\mu}{\omega^2 + \mu^2}; \quad S_{nn}(\omega) = c^2.$$

Предполагается, что сигналы $M(t)$ и $N(t)$ вероятностно независимы:

$$S_{gg}(\omega) = S_{mm}(\omega) + S_{nn}(\omega) = \frac{2D\mu + c^2\omega^2 + \mu^2c^2}{\omega^2 + \mu^2}.$$

Система должна воспроизводить полезный сигнал, т. е.

$$Z(t) = M(t).$$

Тогда

$$K_{zg}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T m(t)[m(t+\tau) + n(t+\tau)]dt = K_{mm}(\tau),$$

так как сигналы $M(t)$ и $N(t)$ независимы и $K_{mn}(\tau) = 0$. Следовательно,

$$S_{zg}(j\omega) = S_{mm}(\omega) = \frac{2D\mu}{\omega^2 + \mu^2}.$$

Произведем факторизацию спектральной плотности $S_{gg}(\omega)$

$$S_{gg}(\omega) = \frac{j\omega + \sqrt{2D\mu + \mu^2c^2}}{j\omega + \mu} \frac{-j\omega + \sqrt{2D\mu + \mu^2c^2}}{-j\omega + \mu}$$

и сепарацию функции $S_{zg}(j\omega)/S_{gg}^-(j\omega)$

$$\frac{S_{zg}(j\omega)}{S_{gg}^-(j\omega)} = \frac{2D\mu}{\omega^2 + \mu^2} \bigg/ \frac{-j\omega + \sqrt{2D\mu + \mu^2c^2}}{-j\omega + \mu} =$$

$$= \frac{2D\mu}{(j\omega + \mu) + (-j\omega + \sqrt{2D\mu + \mu^2c^2})} = 2D\mu \left[\frac{c_1}{j\omega + \mu} + \frac{c_2}{-j\omega + \sqrt{2D\mu + \mu^2c^2}} \right],$$

$$\text{где } c_1 = \frac{1}{\mu c + \sqrt{2D\mu + \mu^2c^2}}, \quad c_2 = \frac{c}{\mu c + \sqrt{2D\mu + \mu^2c^2}}.$$

Подставляя полученные соотношения в (16.14), получаем

$$W_{\text{опт}}(j\omega) = \frac{2D\mu}{\mu c + \sqrt{2D\mu + \mu^2c^2}} \cdot \frac{1}{j\omega + \sqrt{2D\mu + \mu^2c^2}}.$$

Таким образом, для сигналов с заданными спектральными плотностями оптимальная система представляет собой аperiodическое звено с передаточной функцией

$$W_{\text{опт}}(s) = \frac{k_{\text{опт}}}{T_{\text{опт}}s + 1},$$

где коэффициент передачи

$$k_{\text{опт}} = \frac{2D}{2D + \mu c^2 + c\sqrt{2D\mu + \mu^2c^2}} < 1 \quad (16.15)$$

$$T_{\text{опт}} = \frac{c}{\sqrt{2D\mu + \mu^2 c^2}} < \frac{1}{\mu}. \quad (16.16)$$

Отметим, что чем больше интенсивность помехи c^2 и чем шире спектр полезного сигнала (больше μ), тем меньше величина коэффициента передачи $k_{\text{опт}}$ системы.

При отсутствии помехи ($c^2 = 0$) коэффициент передачи $k_{\text{опт}} = 1$, а постоянная времени $T_{\text{опт}} = 0$. Система вырождается в безынерционное звено, которое воспроизводит сигнал без искажений. При отсутствии полезного случайного сигнала ($D = 0$) коэффициент передачи $k_{\text{опт}} = 0$, т. е. система не пропускает ни полезный сигнал, ни помеху.

Необходимо также отметить, что постоянный полезный сигнал оптимальная система будет воспроизводить с ошибкой, поскольку $k_{\text{опт}} < 1$.

Дисперсия ошибки для данной системы

$$\begin{aligned} D_0 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{mm}(\omega) |1 - W(j\omega)|^2 d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{nn}(\omega) |W(j\omega)|^2 d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{2D\mu}{\omega^2 + \mu^2} \left| \frac{Tj\omega + (1-k)}{Tj\omega + 1} \right|^2 d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} c^2 \left| \frac{k}{Tj\omega + 1} \right|^2 d\omega = \\ &= D \frac{\mu T + (1-k)^2}{\mu T + 1} + \frac{c^2 k^2}{2T}. \end{aligned} \quad (16.17)$$

Сравним полученную оптимальную систему (фильтр Винера) с оптимальной системой, имеющей заданную структуру в виде интегрирующего звена, охваченного единичной отрицательной обратной связью, при условии, что на их входы подаются одинаковые полезный сигнал и помеха.

Предположим $D = 1$, $\mu = 1$, $c^2 = 0,5$. Тогда подставляя эти значения в (16.15), (16.16) и (16.17), получаем: $k_{\text{опт}} = 0,55$; $T_{\text{опт}} = 0,45$; $D_0 = 0,61$.

Для системы с заданной структурой в соответствии с (16.2) находим

$$D_0^{\min} = 0,73.$$

Как видим из рассмотренных примеров, применение фиксированной структуры системы может приводить к результату, который отличается от результата, полученного для произвольной структуры.

В том случае, когда разомкнутая часть системы содержит неизменяемую часть, например объект управления, реализация оптимальной передаточной функции возможна за счет введения корректирующих устройств. Например, для последовательной коррекции

$$W_k(s) = \frac{W_{\text{опт}}(s)}{1 - W_{\text{опт}}(s) W_0(s)},$$

где $W_k(s)$ — передаточная функция последовательного корректирующего устройства; $W_0(s)$ — передаточная функция объекта управления; $W_{\text{опт}}(s)$ — оптимальная передаточная функция замкнутой системы.

Однако реализация передаточной функции $W_k(s)$ не всегда возможна. Например, если

$$W_0(s) = \frac{k_0}{s(T_0s + 1)} \text{ и } W_{\text{опт}}(s) = \frac{k_{\text{опт}}}{T_{\text{опт}}s + 1},$$

то передаточная функция устройства коррекции

$$W_k(s) = \frac{k_{\text{опт}}}{k_0} \frac{s(T_0s + 1)}{T_{\text{опт}}s + 1 - k},$$

практическая реализация которой невозможна, так как степень полинома числителя выше степени полинома знаменателя.

16.3.3. Определение передаточной функции системы с ограниченной длительностью переходного процесса (фильтр с конечной памятью)

Выражение (16.14) для оптимальной передаточной функции получено в предположении, что на входе системы действуют только случайные сигналы: полезный и помеха. При этом никаких ограничений на быстроедействие системы не накладывается.

В тех случаях, когда на вход системы помимо случайных сигналов, подается и регулярный, который необходимо воспроизводить, начиная с некоторого момента времени, без ошибки, задача отыскания оптимальной системы становится задачей на условный экстремум.

В новой постановке задачи оптимизации минимум дисперсии ошибки воспроизведения случайного полезного сигнала должен обеспечиваться при соблюдении дополнительного условия, связанного с необходимостью, начиная с некоторого момента времени, точного воспроизведения регулярного сигнала.

Пусть на вход системы подается смесь сигналов

$$G(t) = r(t) + M(t) + N(t),$$

где $r(t)$ — регулярный сигнал, который может быть представлен в виде полинома

$$r(t) = r_0 + r_1 t + r_2 t^2 + \dots + r_l t^l.$$

Для того чтобы переходный процесс в системе заканчивался в течение заданного времени T , необходимо обеспечить условие

$$w(t) \equiv 0 \text{ при } t > T.$$

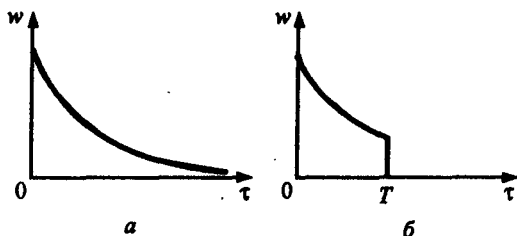


Рис. 16.7. Графики весовых функций фильтров с бесконечной памятью (а) и с конечной памятью (б)

Реакция системы с такой весовой функцией на δ -функцию существует только в течение интервала времени T , т. е. система «запоминает» поданное на него воздействие только в течение ограниченного интервала времени. Вследствие этого она и получила название *системы или фильтра с конечной памятью*.

На рис. 16.7 изображены весовые функции обычного фильтра с бесконечной памятью и фильтра с конечной памятью.

Реакция системы $y_r(t)$ на регулярный сигнал

$$y_r(t) = \int_0^T r(t-\lambda)w(\lambda)d\lambda.$$

Разлагая функцию $r(t-\lambda)$ в ряд Тейлора, ограничиваясь l членами, получаем

$$r(t-\lambda) = r(t) - \lambda \dot{r}(t) + \frac{\lambda^2}{2!} \ddot{r}(t) - \dots + (-1)^l \frac{\lambda^l}{l!} r^{(l)}(t).$$

Тогда

$$y_r(t) = r(t) \int_0^T w(\lambda)d\lambda - \dot{r}(t) \int_0^T \lambda w(\lambda)d\lambda + \dots + (-1)^l \frac{r^{(l)}(t)}{l!} \int_0^T \lambda^l w(\lambda)d\lambda.$$

Ошибка воспроизведения регулярного сигнала:

$$\begin{aligned} \theta_r(t) = r(t) - y_r(t) = r(t) \left[1 - \int_0^T w(\lambda)d\lambda \right] + \\ + \dot{r}(t) \int_0^T \lambda w(\lambda)d\lambda - \dots - (-1)^l \frac{r^{(l)}(t)}{l!} \int_0^T \lambda^l w(\lambda)d\lambda. \end{aligned}$$

Для того чтобы, начиная с момента времени $t = T$, ошибка $\theta_r(t)$ была равна нулю, должны быть равны нулю коэффициенты при всех производных $r^{(i)}(t)$ для $i = 0, 1, \dots, l$:

$$\int_0^T w(\lambda) d\lambda \equiv 1; \quad \int_0^T \lambda w(\lambda) d\lambda \equiv 0; \quad \dots \quad \int_0^T \lambda^l w(\lambda) d\lambda \equiv 0. \quad (16.18)$$

Весовая функция системы, которая обеспечивала бы минимум дисперсии ошибки воспроизведения полезного случайного сигнала на фоне случайной помехи, должна отыскиваться среди весовых функций, которые удовлетворяют условиям (16.18).

Пример 3. Пусть на вход системы поступают только помеха $N(t)$ типа белого шума и скачкообразный регулярный сигнал $r(t) = r_0 1(t)$.

Дисперсия случайного сигнала ошибки, вызванного прохождением помехи на выход, будет равна:

$$D_0^2 = c^2 \int_0^T w^2(\tau) d\tau = J\{w(\tau)\}, \quad (16.19)$$

где c^2 — интенсивность белого шума помехи.

Так как для заданного $r(t)$ все производные равны нулю, ошибка воспроизведения регулярного сигнала, начиная с момента времени T , будет равна нулю лишь при соблюдении условия

$$\int_0^T w(\tau) d\tau \equiv 1.$$

Задача на условный экстремум функционала (16.19) может быть сведена к задаче на безусловный экстремум для функционала

$$I\{w(\tau)\} = J\{w(\tau)\} - 2\lambda_0 \int_0^T w(\tau) d\tau, \quad (16.20)$$

где λ_0 — множитель Лагранжа.

Зададим вариацию функции $w(\tau)$ в виде $\gamma \delta w(\tau)$. Тогда

$$\begin{aligned} I\{w(\tau) + \gamma \delta w(\tau)\} &= c^2 \int_0^T [w(\tau) + \gamma \delta w(\tau)]^2 d\tau - 2\lambda_0 \int_0^T [w(\tau) + \gamma \delta w(\tau)] d\tau = \\ &= c^2 \int_0^T w^2(\tau) d\tau + 2\gamma c^2 \int_0^T \delta w(\tau) w(\tau) d\tau + \gamma^2 c^2 \int_0^T (\delta w(\tau))^2 d\tau - \\ &\quad - 2\lambda_0 \int_0^T w(\tau) d\tau - 2\lambda_0 \gamma \int_0^T \delta w(\tau) d\tau. \end{aligned}$$

Беря производную по параметру γ и приравняв γ к нулю, получаем уравнение для нахождения оптимальной весовой функции:

$$2c^2 \int_0^T \delta w(\tau) w(\tau) d\tau - 2\lambda_0 \int_0^T \delta w(\tau) d\tau = 2 \int_0^T \delta w(\tau) [c^2 w(\tau) - \lambda_0] d\tau = 0.$$

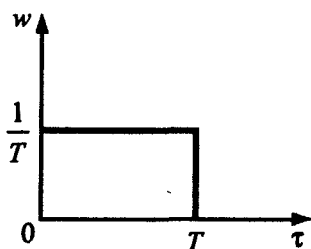


Рис. 16.8. Весовая функция оптимального фильтра с конечной памятью

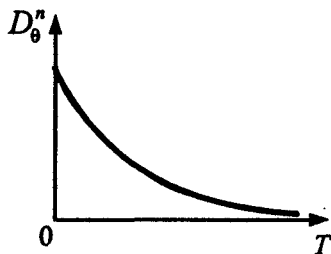


Рис. 16.9. Зависимость дисперсии ошибки от интервала времени

Для выполнения полученного соотношения достаточно соблюдения равенства:

$$c^2 w(\tau) - \lambda_0 = 0.$$

Откуда следует, что оптимальная весовая функция системы удовлетворяет соотношению

$$w(\tau) = \lambda_0 / c^2 = \text{const}$$

при $0 \leq \tau \leq T$. В этом случае

$$\int_0^T w(\tau) d\tau = \lambda_0 T / c^2 = 1,$$

откуда $\lambda_0 = c^2 / T$.

Окончательно оптимальная весовая функция системы имеет вид

$$w(\tau) = 1/T \quad (16.21)$$

при $0 \leq \tau \leq T$. На рис. 16.8 изображен график найденной весовой функции.

Подставляя (16.21) в (16.19), получим

$$D_\theta^n = c^2 / T.$$

Зависимость D_θ^n от интервала времени T представлена на рис. 16.9.

Зависимость D_θ^n от T показывает, что чем выше быстродействие системы, т. е. чем меньше величина T и, следовательно, меньше ее инерционность, тем слабее она подавляет помеху. Эта зависимость характеризует предельные фильтрующие свойства системы при заданном быстродействии и может использоваться для определения минимального значения дисперсии сигнала на выходе системы при действии на ее вход помехи типа белого шума при заданном быстродействии, определяемого величиной T . И наоборот, при заданном уровне дисперсии помехи на выходе с помощью представленной зависимости (см. рис. 16.9) можно определить предельное быстродействие системы.

Передаточная функция, соответствующая оптимальной весовой функции (16.21), имеет вид

$$W_{\text{опт}}(s) = \frac{1 - e^{-sT}}{Ts}.$$

$$W_p(s) = \frac{W_{\text{от}}(s)}{1 - W_{\text{от}}(s)} = \frac{1 - e^{-sT}}{Ts - 1 + e^{-sT}}. \quad (16.22)$$

Реализация (16.22), имеющей вид трансцендентной функции, практически трудно осуществима даже для рассмотренного простого случая.

Если же на вход фильтра поступает помеха, имеющая спектральную плотность $S_{\text{ш}}(\omega) = 2D\mu/(\omega^2 + \mu^2)$, выражение для $w_{\text{от}}(\tau)$ будет содержать слагаемые в виде δ -функций, практическая реализация которых также невозможна.

Рассмотренная постановка задачи представляет интерес главным образом с точки зрения установления предельных возможностей оптимальной фильтрации случайных сигналов с заданными вероятностными характеристиками при заданном быстродействии воспроизведения регулярных сигналов.

16.4. СИНТЕЗ ОПТИМАЛЬНОЙ СИСТЕМЫ В ПРОСТРАНСТВЕ СОСТОЯНИЙ (ФИЛЬТР КАЛМАНА-БЬЮСИ)

Другой подход к оптимальной фильтрации, базирующийся на критерии минимума математического ожидания квадрата случайной ошибки и использующий представление системы в форме пространства состояний, был предложен Р. Калманом и Р. Бьюси (R. Bucy). В отличие от задачи Винера, для задания случайного полезного сигнала в данном случае используется формирующий фильтр, возбуждаемый белым шумом.

Будем рассматривать постановку задачи оптимальной фильтрации, когда предполагается, что поведение системы управления может быть описано системой дифференциальных уравнений в нормальной форме Коши:

$$\begin{aligned} \dot{V}(t) &= AV(t) + BM(t); \\ Y(t) &= CV(t) + N(t), \end{aligned} \quad (16.23)$$

где A , B и C — постоянные матрицы; $V(t)$ — случайный вектор переменных состояния; $Y(t)$ — наблюдаемая случайная переменная выхода; $M(t)$ и $N(t)$ — стационарные случайные воздействия типа идеального белого шума с нулевым математическим ожиданием, причем их корреляционные функции имеют вид:

$$K_{mm}(\tau) = \bar{M}\delta(\tau); K_{nn}(\tau) = \bar{N}\delta(\tau); K_{mn}(\tau) = 0; \bar{M} = \text{const}, \bar{N} = \text{const}.$$

На рис. 16.10 представлена структурная схема, отвечающая матричному описанию (16.23) системы. Формирующий фильтр $\Phi\Phi$, представ-

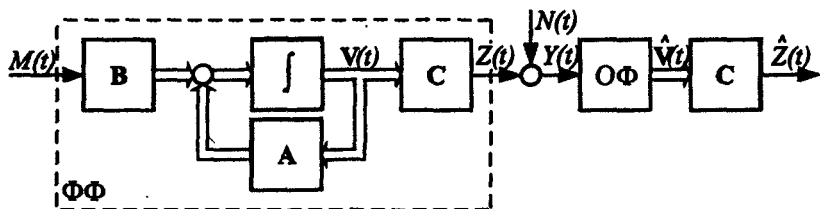


Рис. 16.10. Структурная схема оптимальной системы

ленный на схеме описанием в форме пространства состояний, по заданному воздействию $M(t)$ вырабатывает на выходе требуемый случайный процесс $Z(t) = CV(t)$, на который накладывается помеха $N(t)$.

Оптимальный фильтр ОФ выделяет наилучшим образом в смысле заданного критерия J из наблюдаемого случайного сигнала $Y(t)$ полезный сигнал $Z(t)$ на фоне помехи $N(t)$ в виде оценки $\hat{Z}(t)$.

Синтезируемая оптимальная система (*фильтр Калмана-Бьюси*) должна удовлетворять критерию минимума математического ожидания квадрата случайной ошибки $E(t) = V(t) - \hat{V}(t)$, т. е.

$$J = M[E^2(t)] = M\left[\{V(t) - \hat{V}(t)\}^2\right] = \min. \quad (16.24)$$

Калманом и Бьюси было показано, что оптимальный фильтр, обеспечивающий в любой момент времени воспроизведение полезного сигнала $Z(t)$ при условии (16.24), должен описываться неоднородным дифференциальным уравнением вида

$$\frac{d\hat{V}(t)}{dt} = A\hat{V}(t) + K(t)[Y(t) - \hat{Z}(t)], \quad (16.25)$$

где $\hat{Z}(t) = C\hat{V}(t)$ — оценка выходной величины системы. Дифференциальному уравнению (16.25) соответствует структурная схема оптимального фильтра, приведенная на рис. 16.11.

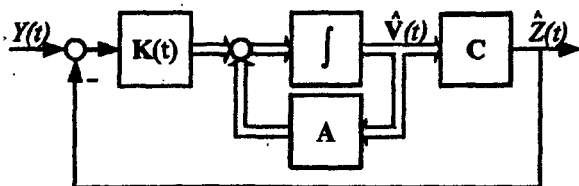


Рис. 16.11. Структурная схема фильтра Калмана-Бьюси

В уравнении (16.25) матрица $K(t)$ является неизвестной и определяется в результате решения системы уравнений:

$$K = PC^T \bar{N}^{-1};$$

$$\frac{dP}{dt} = AP + PA^T - PC^T \bar{N}^{-1} CP + B \bar{M} B^T.$$

Второе уравнение представляет собой дифференциальное уравнение Риккати и решается численно с применением ЭВМ.

В установившемся режиме уравнение Риккати становится алгебраическим:

$$0 = AP + PA^T - PC^T \bar{N}^{-1} CP + B \bar{M} B^T$$

и его решение упрощается.

В скалярном случае, когда $A = a$, $B = C = 1$, $P = p$; $Z = V$, $\hat{Z} = \hat{V}$, искомый коэффициент передачи оптимальной системы определяется решением алгебраической системы уравнений:

$$\frac{1}{\bar{N}} p^2 - 2ap - \bar{M} = 0;$$

$$k = p / \bar{N}.$$

Решая эту систему, получаем

$$k = a + \sqrt{a^2 + \bar{M} / \bar{N}}.$$

При стационарных случайных процессах метод оптимизации Калмана-Бьюси не имеет преимуществ по сравнению с методом оптимизации Винера и дает те же оценки. Однако метод Калмана-Бьюси может быть использован для синтеза оптимальных нестационарных линейных систем.

Глава 17

НЕЛИНЕЙНЫЕ СИСТЕМЫ ПРИ СЛУЧАЙНЫХ ВОЗДЕЙСТВИЯХ

17.1. ПРЕОБРАЗОВАНИЕ СЛУЧАЙНОГО СИГНАЛА НЕЛИНЕЙНЫМ ЗВЕНОМ

Нелинейное преобразование в общем случае вызывает значительное искажение не только детерминированного (см. § 9.2), но и случайного сигнала. Принцип суперпозиции для нелинейных систем неприменим. В связи с этим при воздействии на систему полезного сигнала и

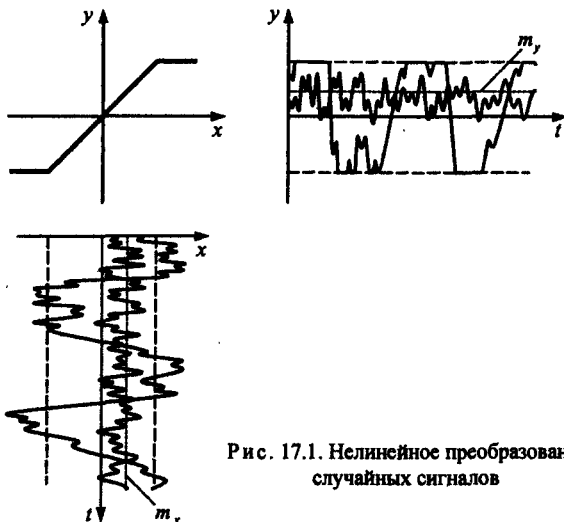


Рис. 17.1. Нелинейное преобразование случайных сигналов

помех, рассмотрение их независимого прохождения через систему невозможно.

Пусть на вход нелинейного элемента (НЭ) приложен стационарный случайный сигнал $X(t)$, который может быть представлен так:

$$X(t) = m_x + \dot{X}(t), \quad (17.1)$$

где $m_x = \text{const}$ — математическое ожидание (среднее значение) входного сигнала, включая регулярную составляющую; $\dot{X}(t)$ — центрированная случайная составляющая входного сигнала.

Рассмотрим прохождение этого сигнала через НЭ на примере безынерционного элемента типа «насыщение» (рис. 17.1).

Если уровень помех невелик, так что входной сигнал остается в пределах линейного участка, то имеет место линейное преобразование:

$$Y(t) = kX(t) = k \left[m_x + \dot{X}(t) \right] = m_y + \dot{Y}(t), \quad (17.2)$$

где $m_y = km_x$ — математическое ожидание выходного сигнала; $\dot{Y}(t) = k\dot{X}(t)$ — центрированная случайная составляющая выходного сигнала, причем каждая из составляющих (17.2) m_y и $\dot{Y}(t)$ зависит от m_x и $\dot{X}(t)$.

Если уровень помех таков, что сигнал выходит за пределы линейного участка характеристики, достигая области насыщения, то математическое ожидание выходного сигнала уменьшается и при дальнейшем росте уровня помех стремится к нулю. Следовательно, большой уро-

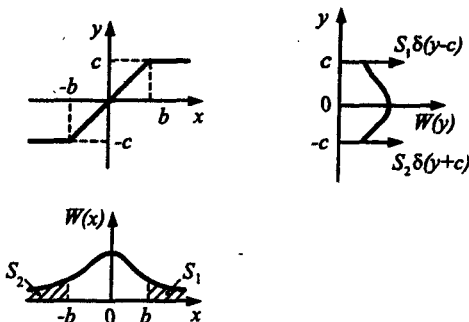


Рис. 17.2. Нелинейное преобразование плотности вероятности сигнала с нормальным распределением

вень помех изменяет полезный сигнал на выходе НЭ, что можно интерпретировать как изменение эквивалентного (усредняющего) коэффициента преобразования НЭ по полезному сигналу. Величина флуктуации сигнала на выходе НЭ ограничена. Спектр случайного сигнала на выходе НЭ обогащается высокочастотными гармониками. Очевидно, что наличие такого элемента в системе, находящейся под действием случайных помех, может существенно повлиять на ее свойства вплоть до потери устойчивости.

Проанализируем влияние НЭ типа «насыщение» на распределение плотности вероятности выходного сигнала (рис. 17.2).

Если входной сигнал $X(t)$ с плотностью вероятности $W(x)$, соответствующей нормальному закону распределения, не выходит за пределы линейного участка характеристики НЭ, то плотность вероятности выходного сигнала $Y(t)$ не меняется, т. е. $W(y) = W(x)$. Поскольку $|y| \leq c$, вероятность появления значения $|y| > c$ равна нулю, т. е. $W(y) = 0$.

Для всех значений входа $x > b$ или $x < -b$ нелинейной характеристики будут соответствовать значения выхода $y = c$ или $y = -c$. Вероятность получения значения $y = c$ ($y = -c$) равна площади заштрихованной области под кривой плотности вероятности $W(x)$, лежащей в интервале $x \in [b, \infty)$ ($x \in (-\infty, -b]$). На графике достижение сигналом уровней насыщения отражается тем, что плотность вероятности $W(y)$ в точках $y = \pm c$ будет представлять собой δ -функции, площади которых равны заштрихованным областям под соответствующими участками кривой плотности вероятности $W(x)$. Окончательно, выражение для плотности вероятности выходного сигнала будет иметь вид

$$W(y) = \begin{cases} W(x) & \text{при } |x| < b; \\ 0 & \text{при } |x| > b; \\ S_1 \delta(y - c) & \text{при } x = b; \\ S_2 \delta(y + c) & \text{при } x = -b. \end{cases}$$

Общая площадь под кривой плотности вероятности выходного сигнала сохраняется равной единице, т. е.

$$\int_{-\infty}^{\infty} W(y) dy = 1.$$

17.2. СТАТИСТИЧЕСКАЯ ЛИНЕАРИЗАЦИЯ НЕЛИНЕЙНОГО ЗВЕНА

Аналитическое исследование случайных процессов в нелинейных системах вызывает значительные трудности вследствие невыполнимости принципа суперпозиции. Нелинейные преобразования, в общем случае, искажают законы распределения вероятностей входных сигналов, меняют их вероятностные характеристики. Поэтому исследования часто ведутся с использованием приближенных методов, базирующихся на статистической линеаризации нелинейностей.

Метод статистической линеаризации, предназначенный для расчета нелинейных систем при случайных процессах, является аналогом метода гармонической линеаризации, применяемого при анализе и синтезе регулярных периодических процессов.

Статистическая линеаризация заключается в замене нелинейной характеристики звена линейной, эквивалентной исходной характеристике звена в отношении вероятностных характеристик сигналов. В результате такой замены возникает возможность использования теории линейных систем с хорошо развитым математическим аппаратом.

С целью упрощения изучения метода и, в то же время, сохраняя его суть, ограничимся рассмотрением прохождения случайного $X(t)$ сигнала (рис. 17.3, а) через НЭ с однозначной нечетной симметричной статической характеристикой вида $y = F(x)$. В соответствии с условиями применения метода статистической линеаризации предполагается, что процесс $X(t)$ на входе звена является стационарным и характеризуется нормальным законом распределения с плотностью вероятности

$$W(x) = \frac{1}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x}. \quad (17.3)$$

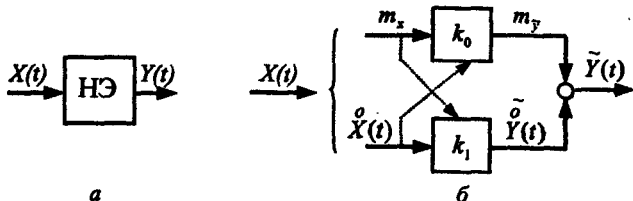


Рис. 17.3. Статистическая линеаризация нелинейного элемента

Представим входной процесс $X(t)$ в виде суммы составляющих (17.1). Установившийся случайный процесс $Y(t)$ на выходе звена, получаемый в результате нелинейного преобразования

$$Y(t) = F[X(t)] \quad (17.4)$$

также может быть представлен суммой составляющих $Y(t) = m_y + \tilde{Y}(t)$.

Задача статистической линеаризации НЭ состоит в замене соотношения (17.4) другим, приближенным (аппроксимирующим) соотношением

$$\tilde{Y}(t) = m_y + \tilde{Y}(t) = k_0 m_x + k_1 \tilde{X}(t), \quad (17.5)$$

где k_0 — коэффициент статистической линеаризации НЭ по регулярной составляющей (математическому ожиданию) процесса; k_1 — коэффициент статистической линеаризации НЭ по случайной составляющей процесса. Коэффициенты k_0 и k_1 подлежат определению. Эквивалентная структура статистически линеаризованного нелинейного звена, соответствующая (17.5), приведена на рис. 17.3, б.

Ввиду того, что полного совпадения сигналов $Y(t)$ и $\tilde{Y}(t)$ в результате статистической линеаризации нелинейного звена достичь невозможно, обычно ограничиваются требованиями совпадения лишь нескольких вероятностных характеристик $Y(t)$ и $\tilde{Y}(t)$. В зависимости от требований (критериев) существуют различные способы нахождения коэффициентов статистической линеаризации. Рассмотрим эти требования.

1. Математические ожидания и дисперсии сигналов $Y(t)$ и $\tilde{Y}(t)$ совпадают, т. е.

$$m_y = m_{\tilde{y}}, \quad (17.6)$$

$$D_y = D_{\tilde{y}}. \quad (17.7)$$

При выполнении условий (17.6) и (17.7) закон распределения выходного сигнала НЭ будет отличаться от соответствующего закона распределения статистически линеаризованного нелинейного звена.

Согласно (17.5) и критерию (17.6), можно записать

$$m_y = k_0 m_x.$$

Следовательно, коэффициент статистической линейаризации по регулярной составляющей процесса

$$k_0 = m_y / m_x. \quad (17.8)$$

В соответствии с (17.5) приравняем случайные составляющие

$$\dot{\tilde{Y}}(t) = k_1 \dot{\tilde{X}}(t). \quad (17.9)$$

Возведя обе части равенства (17.9) в квадрат, получим

$$\left[\dot{\tilde{Y}}(t) \right]^2 = k_1^2 \left[\dot{\tilde{X}}(t) \right]^2. \quad (17.10)$$

Математические ожидания обеих частей (17.10)

$$M \left\{ \left[\dot{\tilde{Y}}(t) \right]^2 \right\} = M \left\{ k_1^2 \left[\dot{\tilde{X}}(t) \right]^2 \right\} = k_1^2 M \left\{ \left[\dot{\tilde{X}}(t) \right]^2 \right\}$$

или, с учетом (17.7),

$$D_y = k_1^2 D_x.$$

Отсюда коэффициент статистической линейаризации по случайной составляющей процесса

$$k_1 = \pm \sqrt{D_y / D_x} = \pm \sigma_y / \sigma_x, \quad (17.11)$$

где $\sigma_y = \sqrt{D_y}$, $\sigma_x = \sqrt{D_x}$ — среднеквадратические отклонения соответственно выходного и входного сигналов НЭ.

Коэффициенты статистической линейаризации k_0 и k_1 однозначно определяются статической характеристикой $F(x)$ НЭ и плотностью распределения входного сигнала.

Математические ожидания в (17.8) вычисляются по формулам:

$$m_x = M[X(t)] = \int_{-\infty}^{\infty} x W(x) dx; \quad m_y = M[Y(t)] = \int_{-\infty}^{\infty} F(x) W(x) dx.$$

Дисперсии процессов в (17.11) на входе и выходе нелинейного элемента:

$$D_x = D[X(t)] = \int_{-\infty}^{\infty} (x - m_x)^2 W(x) dx;$$

$$D_y = D[Y(t)] = \int_{-\infty}^{\infty} (y - m_y)^2 W(y) dy = \int_{-\infty}^{\infty} F^2(x) W(x) dx - m_y^2.$$

Тогда коэффициенты статистической линеаризации по регулярной и случайной составляющим процесса принимают вид:

$$k_0 = \frac{1}{m_x} \int_{-\infty}^{\infty} F(x)W(x)dx; \quad (17.12)$$

$$k_1 = \pm \frac{1}{\sqrt{D_x}} \sqrt{\int_{-\infty}^{\infty} F^2(x)W(x)dx - m_y^2}. \quad (17.13)$$

Знак коэффициента k_1 определяется характером функции $F(x)$. Если функция $F(x)$ возрастает около точки $x = m_x$, то следует взять знак «плюс», если функция $F(x)$ убывает — знак «минус».

2. Среднеквадратическое отклонение разности сигналов $\tilde{Y}(t)$ и $Y(t)$ минимально, т. е.

$$\overline{\varepsilon^2} = M \left\{ [Y(t) - \tilde{Y}(t)]^2 \right\} \rightarrow \min. \quad (17.14)$$

Подставляя (17.5) в выражение (17.14) для среднеквадратического отклонения, получаем

$$\overline{\varepsilon^2} = M \left\{ \left[Y(t) - k_0 m_x - k_1 \dot{X}(t) \right]^2 \right\}.$$

Вычисление среднеквадратического отклонения дает

$$\overline{\varepsilon^2} = D_y + m_y^2 + k_0^2 m_x^2 + k_1^2 D_x - 2k_0 m_x m_y - 2k_1 K_{xy}(0), \quad (17.15)$$

$$\text{где } D_y = M \left\{ \left[\dot{Y}(t) \right]^2 \right\}, \quad D_x = M \left\{ \left[\dot{X}(t) \right]^2 \right\},$$

$K_{xy}(0) = K_{xy}(t, t) = M \left[\dot{X}(t) \dot{Y}(t) \right] = M \left[\dot{X}(t) Y(t) \right]$, причем $K_{xy}(0)$ — значение взаимной корреляционной функции центрированных процессов $\dot{X}(t)$ и $\dot{Y}(t)$.

Для нахождения k_0 и k_1 , обеспечивающих минимум (17.15), необходимо определить частные производные $\overline{\varepsilon^2}$ по каждому из коэффициентов и приравнять их к нулю:

$$\frac{\partial \overline{\epsilon^2}}{\partial k_0} = 2m_x(m_x k_0 - m_y) = 0;$$

$$\frac{\partial \overline{\epsilon^2}}{\partial k_1} = 2[D_x k_1 - K_{xy}(0)] = 0.$$

Следовательно, коэффициенты статистической линеаризации:

$$k_0 = m_y / m_x; \quad (17.16)$$

$$k_1 = K_{xy}(0) / D_x, \quad (17.17)$$

где

$$K_{xy}(0) = M \left[\dot{X}(t)Y(t) \right] = \int_{-\infty}^{\infty} (x - m_x)F(x)W(x)dx.$$

Выражения (17.8) и (17.16) полностью совпадают.

Практически вычисление коэффициента k_1 по формуле (17.13) дает завышенный результат, а по формуле (17.17) — заниженный. Поэтому при расчетах рекомендуется использовать их среднеарифметическое значение

$$k_1 = \frac{k'_1 + k''_1}{2}, \quad (17.18)$$

где k'_1 — значение коэффициента, найденное по (17.13), k''_1 — значение коэффициента, найденное по (17.17).

Рассмотрим примеры определения коэффициентов статистической линеаризации для некоторых типовых НЭ.

Пример 1. Нелинейное звено с характеристикой типа «реле с зоной нечувствительности», на входе которого приложен случайный сигнал с законом распределения (17.3).

Коэффициент статистической линеаризации по регулярной составляющей процесса:

$$\begin{aligned} k_0 &= \frac{1}{m_x} \int_{-\infty}^{\infty} \frac{F(x)}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx = \\ &= \frac{-c}{m_x \sqrt{2\pi D_x}} \int_{-\infty}^{-b} e^{-(x-m_x)^2/2D_x} dx + \frac{c}{m_x \sqrt{2\pi D_x}} \int_b^{\infty} e^{-(x-m_x)^2/2D_x} dx. \end{aligned}$$

Вводя обозначение $u = (x - m_x)/\sqrt{D_x}$, получаем

$$k_0 = \frac{c}{m_x} \left[-\frac{1}{2} + \Phi \left(\frac{m_x + b}{\sqrt{D_x}} \right) + \frac{1}{2} + \Phi \left(\frac{m_x - b}{\sqrt{D_x}} \right) \right] =$$

$$= \frac{c}{m_x} \left[\Phi \left(\frac{m_x + b}{\sqrt{D_x}} \right) + \Phi \left(\frac{m_x - b}{\sqrt{D_x}} \right) \right],$$

где $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-u^2/2} du$ — функция Лапласа (интеграл вероятностей).

Коэффициенты статистической линеаризации по случайной составляющей процесса:

$$k'_1 = \frac{1}{\sqrt{D_x}} \left[\int_{-\infty}^{\infty} \frac{F^2(x)}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx - m_y^2 \right]^{0.5} =$$

$$= \frac{c}{\sqrt{D_x}} \left\{ 1 - \left[\Phi \left(\frac{m_x + b}{\sqrt{D_x}} \right) + \Phi \left(\frac{m_x - b}{\sqrt{D_x}} \right) \right]^2 \right\}^{0.5};$$

$$k''_1 = \frac{1}{D_x} \left[-c \int_{-\infty}^{-b} \frac{x-m_x}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx + c \int_b^{\infty} \frac{x-m_x}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx \right] =$$

$$= \frac{c}{\sqrt{2\pi D_x}} \left[e^{-(b+m_x)^2/2D_x} + e^{-(b-m_x)^2/2D_x} \right].$$

В частном случае, когда $b=0$, характеристика НЭ вырождается в характеристику типа «идеальное реле» и коэффициенты приобретают вид:

$$k_0 = \frac{2c}{m_x} \Phi \left(\frac{m_x}{\sqrt{D_x}} \right);$$

$$k'_1 = \frac{c}{\sqrt{D_x}} \sqrt{1 - 4 \left[\Phi \left(\frac{m_x}{\sqrt{D_x}} \right) \right]^2}; \quad k''_1 = \frac{2c}{\sqrt{2\pi D_x}} e^{-m_x^2/2D_x}.$$

На рис. 17.4 приведены графики зависимостей коэффициентов статистической линеаризации идеального реле от характеристик m_x и D_x [73]. Кривая 1 (рис. 17.4, б) соответствует коэффициенту k'_1 , кривая 2 — k''_1 . Из графиков видно, что при малых m_x НЭ с характеристикой типа «идеальное реле» ведет себя как линейный усилитель. При больших m_x коэффициент передачи по случайной составляющей процесса ослабляется за счет насыщения НЭ.

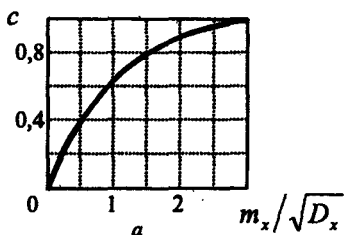
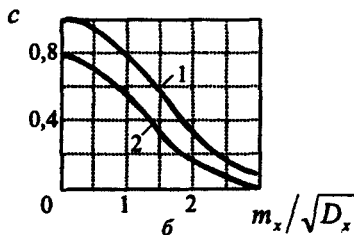
$\frac{m_x k_0}{c}$  $\sqrt{D_x} k_1$ 

Рис. 17.4. Графики зависимости коэффициентов статистической линейизации идеального реле от характеристик m_x и D_x

Пример 2. Нелинейное звено с характеристикой типа «усилитель с насыщением». Коэффициент статистической линейизации по регулярной составляющей процесса:

$$\begin{aligned}
 k_0 &= \frac{1}{m_x} \int_{-\infty}^{\infty} \frac{F(x)}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx = \\
 &= \frac{1}{m_x} \left[-c \int_{-\infty}^{-b} \frac{1}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx + \frac{c}{b} \int_{-b}^b \frac{x}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx + \right. \\
 &\left. + c \int_b^{\infty} \frac{1}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx \right] = \frac{c\sqrt{D_x}}{bm_x\sqrt{2\pi}} \left(e^{-(b+m_x)^2/2D_x} - e^{-(b-m_x)^2/2D_x} \right) + \\
 &+ c \left(\frac{1}{m_x} + \frac{1}{b} \right) \Phi \left(\frac{m_x+b}{\sqrt{D_x}} \right) + c \left(\frac{1}{m_x} - \frac{1}{b} \right) \Phi \left(\frac{m_x-b}{\sqrt{D_x}} \right).
 \end{aligned}$$

Коэффициенты статистической линейизации при случайной составляющей процесса:

$$\begin{aligned}
 k_1' &= \frac{1}{\sqrt{D_x}} \left[\int_{-\infty}^{\infty} \frac{F^2(x)}{\sqrt{2\pi D_x}} e^{-(x-m_x)^2/2D_x} dx - k_0^2 m_x^2 \right]^{0.5} = \\
 &= \frac{c}{\sqrt{D_x}} \left\{ 1 - \left(\frac{m_x^2}{b^2} + \frac{D_x}{b^2} \right) \left[\Phi \left(\frac{b+m_x}{\sqrt{D_x}} \right) + \Phi \left(\frac{b-m_x}{\sqrt{D_x}} \right) \right] - \right. \\
 &\left. - \frac{\sqrt{D_x}}{b\sqrt{2\pi}} \left[\left(1 + \frac{m_x}{b} \right) e^{-(b-m_x)^2/2D_x} + \left(1 - \frac{m_x}{b} \right) e^{-(b+m_x)^2/2D_x} - k_0^2 m_x^2 \right]^{0.5} \right\};
 \end{aligned}$$

$$k_1^* = \frac{c}{\sqrt{2\pi D_x}} \left[e^{-(b+m_x)^2/2D_x} + e^{-(b-m_x)^2/2D_x} \right] + \frac{c}{b\sqrt{2\pi}} \frac{\frac{b-m_x}{\sqrt{D_x}}}{\frac{-(b+m_x)}{\sqrt{D_x}}} \int_{\frac{-(b+m_x)}{\sqrt{D_x}}}^{\frac{b-m_x}{\sqrt{D_x}}} u^2 e^{-u^2/2} du +$$

$$+ \frac{c}{\sqrt{2\pi D_x}} \frac{m_x}{b} \left[e^{-(b+m_x)^2/2D_x} + e^{-(b-m_x)^2/2D_x} \right].$$

При выводе использовались соотношения:

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du = \Phi(b) - \Phi(a);$$

$$\frac{1}{\sqrt{2\pi}} \int_a^b u e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \left(e^{-a^2/2} - e^{-b^2/2} \right);$$

$$\frac{1}{\sqrt{2\pi}} \int_a^b u^2 e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \left(a e^{-a^2/2} - b e^{-b^2/2} \right) + \Phi(b) - \Phi(a).$$

17.3. ИССЛЕДОВАНИЕ НЕЛИНЕЙНЫХ СИСТЕМ МЕТОДОМ СТАТИСТИЧЕСКОЙ ЛИНЕАРИЗАЦИИ

Рассмотрим применение метода статистической линеаризации для расчета замкнутых систем. Структурная схема (рис. 17.5, а) содержит один безынерционный НЭ с однозначной статической характеристикой $F(x)$ и динамическую линейную часть (ЛЧ) с передаточной функцией $W(s)$, включающую инерционные звенья. Предполагается также, что входной сигнал $G(t)$ системы, который в общем случае представляет собой аддитивную смесь полезного сигнала и помехи, является стационарным случайным процессом с нормальным законом распределения. Входной сигнал может быть представлен в виде

$$G(t) = m_g + \dot{G}(t),$$

где m_g — регулярная составляющая $G(t)$; $\dot{G}(t)$ — случайная центрированная составляющая.

В соответствии с методом статистической линеаризации необходимо, чтобы сигнал $X(t)$ на входе НЭ подчинялся нормальному закону распределения, т. е. выполнялась гипотеза нормального закона распределения сигнала. Последнее возможно лишь в том случае, когда выпол-

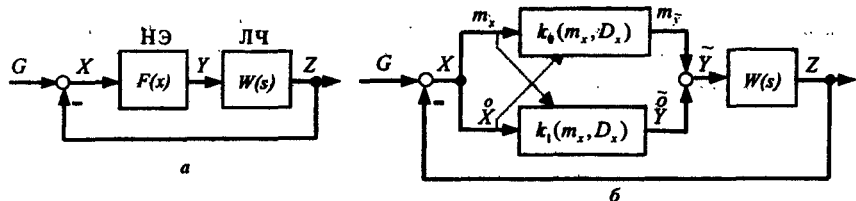


Рис. 17.5. Нелинейная система (а)
и ее эквивалентная статистически линеаризованная модель (б)

няется гипотеза фильтра линейной части системы. Смысл гипотезы фильтра состоит в следующем.

Закон распределения случайного сигнала $Y(t)$ на выходе НЭ, как правило, отличается от нормального. Однако, пройдя через ЛЧ системы, обладающей свойством низкочастотного фильтра, сигнал на выходе системы нормализуется. В результате сигнал на входе НЭ, представляющего собой алгебраическую сумму нормально распределенных сигналов, также будет иметь нормальный закон распределения.

Таким образом, здесь имеет место аналогия с гипотезой формы сигнала и гипотезой фильтра ЛЧ, применявшихся в методе гармонической линеаризации при исследовании периодических сигналов.

Используя метод статистической линеаризации по отношению к НЭ, можно перейти к рассмотрению эквивалентной структурной схемы (рис. 17.5, б). Нелинейному элементу соответствуют два канала с коэффициентами передачи k_0 и k_1 , зависящих от вероятностных характеристик сигнала X . По одному каналу проходит регулярная составляющая, по другому — случайная составляющая.

Эквивалентная статистически линеаризованная модель распадается на две: одну, составленную относительно регулярных сигналов (рис. 17.6, а), другую — относительно центрированных случайных сигналов (рис. 17.6, б).

Передаточные функции замкнутой статистически линеаризованной системы относительно ошибки X :

- по регулярной составляющей

$$\Phi_{m_x}(s) = \frac{1}{1 + k_0(m_x, D_x)W(s)}; \quad (17.19)$$

- по центрированной случайной составляющей

$$\Phi_{\tilde{x}}(s) = \frac{1}{1 + k_1(m_x, D_x)W(s)}. \quad (17.20)$$

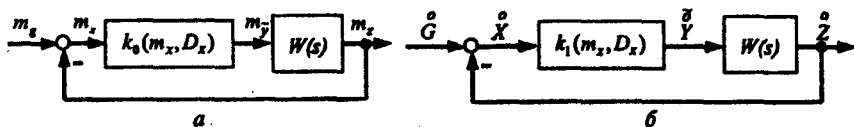


Рис. 17.6. Статистически линеаризованные модели при регулярных (а) и центрированных случайных (б) сигналах

Взаимосвязь между передаточными функциями (17.19) и (17.20) осуществляется через коэффициенты статистической линеаризации k_0 и k_1 , которые являются функциями искомых величин m_x и D_x .

Изображение по Лапласу математического ожидания ошибки в соответствии с моделью, представленной на рис. 17.6, а, имеет вид

$$L\{m_x(t)\} = \Phi_{m_x}(s) L\{m_g(t)\}.$$

В связи с тем, что рассматриваются установившиеся стационарные процессы, получаем

$$m_x = \text{const}, \quad m_g = \text{const}.$$

Тогда можно записать

$$m_x = \frac{m_g}{1 + k_0(m_x, D_x)W(0)}. \quad (17.21)$$

В том случае, если система астатическая, т. е. коэффициент усиления линейной части $W(0)$ на нулевой частоте равен бесконечности:

$$m_x = m_y = 0; \quad m_z = m_g.$$

Дисперсия случайной составляющей ошибки, используя модель, приведенную на рис. 17.6, б, определяется так:

$$D_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \Phi_{\dot{x}}(j\omega) \right|^2 S_{gg}(\omega) d\omega, \quad (17.22)$$

где

$$\left| \Phi_{\dot{x}}(j\omega) \right|^2 = \left| \frac{1}{1 + k_1(m_x, D_x)W(j\omega)} \right|^2.$$

$S_{xx}(\omega)$, $S_{gg}(\omega)$ — спектральные плотности центрированных случайных составляющих ошибки и входного сигнала соответственно.

С учетом выражений (17.21), (17.22) можно записать систему из двух юнечных нелинейных уравнений:

$$\begin{aligned} m_x &= f_1[k_0(m_x, D_x)]; \\ D_x &= f_2[k_1(m_x, D_x)]. \end{aligned} \quad (17.23)$$

Система уравнений (17.23) содержит четыре неизвестных величины m_x , D_x , k_0 , k_1 . Для ее решения необходимо записать еще два уравнения, которые получаются из одного из ранее рассмотренных способов статистической линеаризации (17.8), (17.11) или (17.16), (17.17):

$$\begin{aligned} k_0 &= f_3(m_x, D_x); \\ k_1 &= f_4(m_x, D_x). \end{aligned} \quad (17.24)$$

В общем случае решение системы конечных уравнений (17.23), (17.24) возможно одним из *численных методов*. В частности, эффективным численным методом является метод Ньютона (I. Newton) с использованием принципа продолжения по параметру.

Для решения системы уравнений (17.23), (17.24) может использоваться также *графоаналитический метод*, который состоит в следующем.

1. В системе координат (m_x, f_1) строится семейство кривых, соответствующих первому уравнению (17.23) при различных значениях дисперсии $D_{x_i} = \text{const}$ (рис. 17.7, а).

2. Из начала координат проводится биссектриса угла и по точкам пересечения ее с кривыми семейства $f_1(m_x)$ строится график зависимости $D_x = \varphi_1(m_x)$ (рис. 17.7, в).

3. В системе координат (D_x, f_2) строится семейство кривых, соответствующих второму уравнению (17.23) при различных значениях $m_{x_i} = \text{const}$ (рис. 17.7, б).

4. Из начала координат проводится биссектриса угла и по точкам пересечения ее с кривыми семейства $f_2(D_x)$ строится график зависимости $m_x = \varphi_2(D_x)$ (рис. 17.7, в).

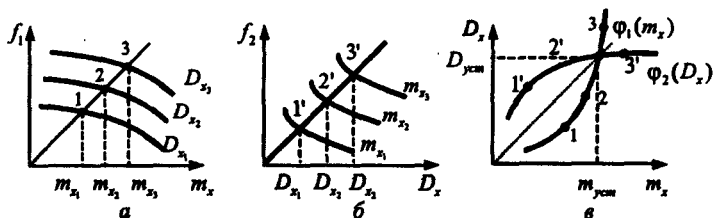


Рис. 17.7. Графоаналитический способ определения вероятностных характеристик замкнутой нелинейной системы

Точки пересечения кривых $\varphi_1(m_x)$ и $\varphi_2(D_x)$ представляют собой решения системы (17.23) и определяют значения $m_{уст}$ и $D_{уст}$ в установившемся режиме.

Метод статистической линеаризации вследствие своей простоты и достаточной для практики точности нашел широкое применение в инженерных расчетах нелинейных систем высокого порядка.

Пример 3. Определить математическое ожидание m_x и дисперсию D_x установившейся ошибки системы (см. рис. 17.5, а) с НЭ типа идеального реле $y = c \operatorname{sgn} x$ и ЛЧ с передаточной функцией $W(s) = k/s$. На входе системы приложен стационарный случайный сигнал $G(t)$ с характеристиками $m_g = \text{const}$, $S_{gg}(\omega) = 2D\mu/(\mu^2 + \omega^2)$.

В силу того, что система обладает астатизмом, математическое ожидание установившейся ошибки $m_x = 0$.

Передаточная функция статистически линеаризованной системы по ошибке относительно случайной составляющей:

$$\Phi_{\circ x}(s) = \frac{s}{s + k_1 k}.$$

Спектральная плотность сигнала ошибки

$$\begin{aligned} S_{xx}(\omega) &= \left| \Phi_{\circ x}(j\omega) \right|^2 S_{gg}(\omega) = \\ &= \frac{-(j\omega)^2}{|j\omega + k_1 k|^2} \frac{2D\mu}{\mu^2 + \omega^2} = \frac{-2D\mu(j\omega)^2}{|(j\omega)^2 + (\mu + k_1 k)(j\omega) + \mu k_1 k|^2}. \end{aligned}$$

Дисперсия сигнала ошибки

$$\begin{aligned} D_x &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-2D\mu(j\omega)^2}{|(j\omega)^2 + (\mu + k_1 k)(j\omega) + \mu k_1 k|^2} d\omega = \\ &= \frac{2D\mu^2 k_1 k}{2\mu k_1 k(\mu + k_1 k)} = \frac{D\mu}{\mu + k_1 k}. \end{aligned} \quad (17.25)$$

Учитывая, что функция Лапласа $\Phi(0) = 0$, коэффициенты статистической линеаризации релейного элемента:

$$k'_1 = \frac{c}{\sqrt{D_x}}; \quad k''_1 = \frac{2c}{\sqrt{2\pi D_x}}.$$

Тогда используя (17.18), усредненный коэффициент статистической линеаризации

$$k_1 = c\beta/\sqrt{D_x}, \quad (17.26)$$

где $\beta = 0,5 + 1/\sqrt{2\pi}$.

Рассматривая совместно (17.25) и (17.26), получим

$$\frac{c^2 \beta^2}{k_1^2} = \frac{D\mu}{\mu + k_1 k},$$

откуда

$$D\mu k_1^2 - c^2 \beta^2 k k_1 - c^2 \beta^2 \mu = 0.$$

Следовательно,

$$k_1 = \frac{c^2 \beta^2 k + \sqrt{c^4 \beta^4 k^2 + 4c^2 \beta^2 D\mu^2}}{2D\mu}.$$

Подставляя k_1 в (17.25), окончательно получаем

$$D_x = \frac{2D^2 \mu^2}{2D\mu^2 + kc\beta \left(kc\beta + \sqrt{k^2 c^2 \beta^2 + 4D\mu^2} \right)}.$$

ЧАСТЬ 6

ОПТИМАЛЬНЫЕ СИСТЕМЫ УПРАВЛЕНИЯ

Глава 18

УСЛОВИЯ ОПТИМАЛЬНОСТИ ПРОЦЕССОВ В ДИНАМИЧЕСКИХ СИСТЕМАХ

18.1. ПРИНЦИП МАКСИМУМА

При решении задач оптимального управления в настоящее время широко применяется метод, получивший название принцип максимума, разработанный Л.С. Понтрягиным и его коллегами в 1956–1960 гг. [49]. С помощью принципа максимума удалось не только обобщить достижения классического вариационного исчисления, полученные Эйлером, Лагранжем и др., но и дать мощный инструмент для решения так называемых неклассических задач оптимизации, в которых требуется учесть ограничения на переменные состояния и управляющие координаты. При этом управление $u(t)$ отыскивается в более широком классе функций, а именно кусочно-непрерывных с разрывами первого рода в отдельные моменты времени. В задачах с переменными, изменяющимися в открытой области, условия экстремума в форме принципа максимума и классического вариационного исчисления совпадают. Поэтому ограничимся рассмотрением только принципа максимума. С методами классического вариационного исчисления можно познакомиться, например, в [32].

18.1.1. Постановка задачи

Рассмотрим стационарный объект управления, описываемый дифференциальными уравнениями

$$dx_i / dt = f_i(x_1, \dots, x_n, u_1, \dots, u_m), \quad i = \overline{1, n}, \quad (18.1)$$

или в векторной форме

$$dx/dt = f(x, u),$$

где функции $f(x, u)$ непрерывны по x, u , имеют непрерывные производные по x . В реальных условиях изменение управления ограничено некоторой заданной допустимой областью

$$u(t) \in U, \quad (18.2)$$

например в виде ограничений на диапазон изменения

$$|u_i| \leq \bar{u}_i, \quad i = \overline{1, m},$$

при этом допустимое множество U есть m -мерный параллелепипед. Назовем управление допустимым, если оно является кусочно-непрерывной функцией при $0 < t < t_k$, непрерывным при $t = 0$ и $t = t_k$, и принадлежит допустимой области (18.2). Анализ большого числа систем управления показывает, что по характеру изменения во времени оптимальное управление следует искать среди кусочно-непрерывных функций, в частности кусочно-постоянных. Так, установлено без применения вариационных методов, что минимальное время перехода объекта из одного состояния в другое достигается при релейном управлении $u(t)$, принимающем два значения u или $-u$.

Пусть задан показатель качества в виде интегрального функционала с функцией $f_0(\cdot)$, не зависящей от времени t ,

$$J = \int_0^{t_k} f_0(x, u) dt \rightarrow \min_{u \in U}. \quad (18.3)$$

В принципе максимума осуществляется переход от интегрального функционала (18.3) к неинтегральному (задача Майера). С этой целью введем дополнительную переменную $x_0(t)$, удовлетворяющую уравнению

$$dx_0/dt = f_0(x, u), \quad x_0(0) = 0. \quad (18.4)$$

Тогда $x_0(t_k) = J$, т.е. равно значению показателя качества (18.3). Присоединив уравнение (18.4) к уравнению объекта (18.1), получим расширенную систему дифференциальных уравнений

$$\begin{aligned} \dot{x}_0 &= f_0(x, u), \\ \dot{x}_1 &= f_1(x, u), \\ &\dots \dots \dots, \\ \dot{x}_n &= f_n(x, u) \end{aligned} \quad (18.5)$$

или в векторном виде $\dot{\tilde{x}} = \tilde{f}(x, u)$,

где $\tilde{x}^T = [x_0, x_1, \dots, x_n]$, $\tilde{f}^T = [f_0, f_1, \dots, f_n]$. Отметим, что правые части уравнений (18.5) не зависят от x_0 .

Приведем геометрическую иллюстрацию для $n = 2$. С учетом координаты x_0 рассмотрим трехмерное пространство (рис. 18.1). Движение начинается из точки A_0 с координатами $x(0) = x_0, x_0(0) = 0$ и заканчи-

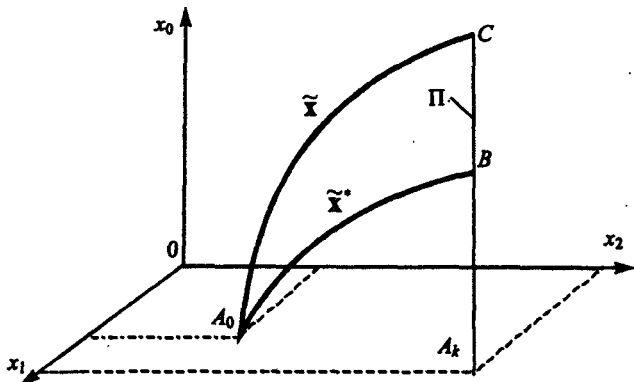


Рис. 18.1. Оптимальная траектория в расширенном пространстве

вадается в точке A_k , $\mathbf{x}(t_k) = \mathbf{x}_k$, на плоскости x_1, x_2 . Из точки A_k построим вертикальный луч Π , параллельный оси x_0 . Любая траектория $\mathbf{X}(t)$ при допустимом управлении $\mathbf{u}(t) \in U$, обеспечивающая переход в заданное конечное состояние $\mathbf{x}(t_k) = \mathbf{x}_k$, заканчивается на луче Π (траектории A_0B , A_0C). При оптимальном управлении отрезок, отсекаемый на луче Π , имеет наименьшую длину, например A_kB .

18.1.2. Формулировка принципа максимума

Введем вспомогательные переменные $\psi_i(t)$, $i = \overline{0, n}$, удовлетворяющие системе дифференциальных уравнений

$$d\psi_i/dt = - \sum_{j=0}^n \frac{\partial f_j(\mathbf{x}, \mathbf{u})}{\partial x_i} \psi_j, \quad i = \overline{0, n} \quad (18.6)$$

или в векторной форме

$$d\tilde{\psi}/dt = - \left[\frac{\partial f(\mathbf{x}, \mathbf{u})}{\partial \tilde{\mathbf{x}}} \right]^T \tilde{\psi}. \quad (18.6a)$$

Переменные $\psi_i(t)$, $i = \overline{0, n}$ непрерывны, имеют непрерывные производные по t всюду кроме конечного числа точек, соответствующих точкам разрыва $\mathbf{u}(t)$.

Запишем функцию Гамильтона (гамильтониан)

$$\tilde{H}(\tilde{\psi}, \mathbf{x}, \mathbf{u}) = \sum_{j=0}^n \psi_j f_j(\mathbf{x}, \mathbf{u}) = \tilde{\psi}^T \tilde{\mathbf{f}} = \tilde{\mathbf{f}}^T \tilde{\psi}. \quad (18.7)$$

Из (18.7) следует, что гамильтониан является скалярной функцией $2n + 1 + m$ переменных.

В результате анализа системы (18.6) и функции (18.7) приведем уравнения, более удобные при формировании правых частей системы (18.6),

$$d\psi_i / dt = -\partial \tilde{H} / \partial x_i, \quad i = \overline{0, n} \quad \text{или} \quad d\tilde{\psi} / dt = -\partial \tilde{H} / \partial \tilde{x}. \quad (18.8)$$

Кроме того

$$dx_i / dt = \partial \tilde{H} / \partial \psi_i, \quad i = \overline{0, n} \quad \text{или} \quad d\tilde{x} / dt = \partial \tilde{H} / \partial \tilde{\psi}. \quad (18.9)$$

Уравнения (18.8) и (18.9), рассматриваемые совместно, называются каноническими уравнениями Гамильтона. При фиксированных $\tilde{x}(t)$ и $\tilde{\psi}(t)$ гамильтониан $\tilde{H}$ есть функция только переменной u .

Пусть максимальное значение $\tilde{H}$ достигается при $u(t) = u^*(t) \in U$

$$\max_{u \in U} \tilde{H}(\tilde{\psi}, x, u) = \tilde{H}(\tilde{\psi}, x, u^*).$$

Приведем формулировку принципа максимума для объекта (18.1), критерия (18.3) с закрепленными концами траектории $x(0) = x_0$, $x(t_k) = x_k$ и свободным временем перехода t_k .

Для оптимальности $u^*(t)$ траектории $x^*(t)$ необходимо существование непрерывной ненулевой вектор-функции $\tilde{\psi}^*(t)$, являющейся решением уравнения (18.6) и такой, что:

- функция $\tilde{H}(\tilde{\psi}, x, u)$ переменной $u \in U$ при фиксированных $\tilde{\psi}^*$, x^* достигает максимума при $u(t) = u^*(t)$, т.е. $\max_{u \in U} \tilde{H}(\tilde{\psi}^*, x^*, u) = \tilde{H}(\tilde{\psi}^*, x^*, u^*)$;

- в конечный момент времени t_k выполняются соотношения

$$\psi_0(t_k) \leq 0, \quad (18.10)$$

$$\tilde{H}(\tilde{\psi}^*(t_k), x^*(t_k), u^*(t_k)) = 0. \quad (18.11)$$

При этом оказывается, что переменная $\psi_0(t) = \text{const}$ для любого $t \in [0, t_k]$, а также

$$\max_{u \in U} \tilde{H}(\tilde{\psi}^*(t), x^*(t), u(t)) = \tilde{H}(\tilde{\psi}^*(t), x^*(t), u^*) = 0. \quad (18.12)$$

Убедимся в справедливости этого соотношения. Вычислим производную

$$\frac{d\tilde{H}}{dt} = \left(\frac{\partial \tilde{H}}{\partial u} \right)^T \frac{du}{dt} + \left(\frac{\partial \tilde{H}}{\partial \tilde{\psi}} \right)^T \frac{d\tilde{\psi}}{dt} + \left(\frac{\partial \tilde{H}}{\partial \tilde{x}} \right)^T \frac{d\tilde{x}}{dt}.$$

Учитывая уравнения Гамильтона (18.8), (18.9), получим

$$\frac{d\tilde{H}}{dt} = \left(\frac{\partial \tilde{H}}{\partial u} \right)^T \frac{du}{dt}.$$

Анализ этого выражения показывает, что $d\tilde{H}/dt = 0$. Действительно, если максимум $\tilde{H}(\tilde{\psi}, x, u)$ достигается во внутренней точке $u(t) \in U$, то

$\partial \tilde{H} / \partial u = 0$ в этой точке. Если $\tilde{H}(\tilde{\psi}, x, u)$ имеет максимум на границе множества U , то $\partial \tilde{H} / \partial u \neq 0$, но в этом случае оказывается, что векторы $\partial \tilde{H} / \partial u$ и du/dt взаимно ортогональны, и следовательно, $d\tilde{H}/dt = 0$. В частном случае, когда $u = \text{const}$ на границе U , $du/dt = 0$ и значит, $d\tilde{H}/dt = 0$. Равенство $d\tilde{H}/dt = 0$ совместно с условием (18.11) приводит к соотношению (18.12).

Запишем уравнение и его решение для переменной $\psi_0(t)$

$$\dot{\psi}_0 = -\partial \tilde{H} / \partial x_0 = 0, \quad \psi_0(t) = \text{const}.$$

Так как функция $\tilde{H}$ (18.7) однородна относительно переменных $\psi_0, \psi_1, \dots, \psi_n$, то условие ее максимума не изменится при делении на константу $\psi_0 \neq 0$. В принципе максимума принято, что $\psi_0 = -1$. Случай $\psi_0 = 0$ относится к *особому* и в дальнейшем не рассматривается.

Отметим, что иногда применяется *принцип минимума*, в котором приняты вспомогательные переменные $p_0 = 1, p_i(t) = -\psi_i(t), i = 1, n$, [10].

Заметим, что принцип максимума есть необходимое условие экстремума. Это означает, что если существует оптимальное решение, то оно удовлетворяет принципу максимума. С другой стороны, решение, полученное на его основе, не обязательно является оптимальным, а будет лишь «подозрительным» на оптимальность. Практически оптимальность устанавливается путем анализа решений, в том числе на ЭВМ, а также используя опыт разработчика.

18.1.3. Применение принципа максимума

При решении задачи оптимизации сначала из условия максимума функции $\tilde{H}$ по $u \in U$ определяют оптимальное управление как функцию x и ψ , т.е. $u = \Phi(x, \psi)$. После подстановки в уравнения (18.1) и (18.8) получим систему из $2n$ уравнений

$$\dot{x} = f(x, \Phi(x, \psi)), \quad \dot{\psi} = -\partial \tilde{H}(x, \psi, \Phi(x, \psi)) / \partial x. \quad (18.13)$$

Определение оптимальных решений $u^*(t), x^*(t)$ при этом сводится к решению *краевой* или *двухточечной* граничной задачи для уравнений (18.13). В отличие от задачи Коши, в которой эти уравнения решают при известных $2n$ начальных условиях $x(0), \psi(0)$, в краевой задаче n условий даны в начальной точке $x(0) = x_0$, а другие n условий — в конечной точке $x(t_k) = x_k$.

В задаче со свободным временем t_k используется дополнительно условие (18.12). Таким образом, имеем полный набор $2n + 1$ условий, необходимый для решения краевой задачи. Если t_k фиксировано, то

выполнения условия (18.12) не требуется, так как в этом случае число условий равно $2n + 1$.

На практике решение во многих случаях связано с большим объемом вычислений, причем решение в аналитическом виде часто невозможно, и поэтому все большее применение находят различные численные методы решения краевой задачи.

В том случае, когда гамильтониан линейно зависит от u , определение $u = \varphi(x, \psi)$ в процессе максимизации функции $\tilde{H}$ по $u \in U$ невозможно. В этом случае нахождение конкретного вида управления $u_i(t)$ связано с применением дополнительного необходимого условия оптимальности

$$(-1)^k \frac{\partial}{\partial u_i} \left[\frac{d^{2k}}{dt^{2k}} \frac{\partial \tilde{H}}{\partial u_i} \right] \leq 0, \quad k = 1, 2, \dots, \quad i = \overline{1, m}.$$

18.1.4. Условия трансверсальности

Принцип максимума был сформулирован в п. 18.1.2 для системы с заданными начальной $x(0) = x_0$ и конечной $x(t_k) = x_k$ точками. Рассмотрим задачу с подвижными концами, когда начальная и (или) конечная точки должны располагаться на некоторых заранее заданных поверхностях.

Пусть в пространстве $X \subset R^n$ заданы k гиперповерхностей, $k \leq n - 1$,

$$h_i(x) = 0, \quad i = \overline{1, k}. \quad (18.14)$$

Для определенности примем, что начальная точка $x(0)$ задана, а конечная точка $x(t_k)$ должна принадлежать пересечению гиперповерхностей (18.14). Множество всех векторов x , удовлетворяющих одновременно условиям (18.14), называется $(n - k)$ -мерным многообразием в пространстве $X \subset R^n$ при выполнении условия

$$\text{rank } d h / d x = k, \quad (18.15)$$

что означает линейную независимость векторов-градиентов $\partial h_i / \partial x = \text{grad } h_i(x)$, $i = \overline{1, k}$. Обозначим $(n - k)$ -мерное многообразие, соответствующее условиям (18.14), через $P(x)$ (если $n - k = 1$, то $P(x)$ есть линия; например, при $n = 3$, $k = 2$ пересечение двух поверхностей $h_1(x) = 0$, $h_2(x) = 0$ дает линию в трехмерном пространстве). Если предположить, что конечная точка известна, то придем к задаче с фиксированными концами. Поэтому необходимые условия оптимальности в виде принципа максимума остаются справедливыми и для задачи с подвижными концами.

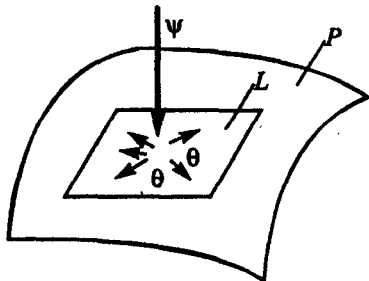


Рис. 18.2. Геометрическая иллюстрация условия трансверсальности

Как отмечалось в § 18.1, для решения краевой задачи требуется знать $2n$ условий (не включая конечный момент времени t_k). Из исходных данных имеем $(n + k)$ условий, т.е. требуется определить недостающие $(n - k)$ соотношений, которые называются условиями трансверсальности.

Условие трансверсальности состоит в том, чтобы вектор $\psi(t_k)$ в искомой точке $x(t_k)$ был ортогонален гиперплоскости L , касательной к многообразию P , рис. 18. 2. Это означает, что $\psi(t_k)$ должен быть ортогонален любому вектору $\theta \in L$:

$$\psi^T \theta \Big|_{t=t_k} = \sum_{i=1}^n \psi_i \theta_i \Big|_{t=t_k} = 0. \quad (18.16)$$

Для определения θ запишем уравнение касательной гиперплоскости L , к гиперповерхности $h_i(x) = 0$, $i = \overline{1, k}$:

$$\sum_{j=1}^n \frac{\partial h_i(x)}{\partial x_j} \theta_j = 0, \quad i = \overline{1, k}. \quad (18.17)$$

Условие трансверсальности (18.16), рассматриваемое совместно с уравнениями (18.17) и условием (18.15), позволяет найти недостающие $(n - k)$ соотношений для вспомогательных переменных $\psi_1(t_k), \dots, \psi_n(t_k)$. Аналогичные условия используются в случае левого подвижного конца траектории при $t = 0$.

Рассмотрим частные практически важные случаи задания исходных данных. Пусть задана начальная точка $x(0) = x_0$, время t_k , а координаты $x_1(t_k), \dots, x_n(t_k)$ — свободны. Для таких условий, наложенных на правый конец траектории, многообразие $P(x)$ совпадает со всем пространством R^n . Следовательно, вектор $\psi(t_k)$ должен быть ортогонален любому вектору в R^n . Поэтому условие трансверсальности (18.16) может быть выполнено только при

$$\psi_1(t_k) = 0, \dots, \psi_n(t_k) = 0, \quad (18.18)$$

что дает недостающие n условий для решения краевой задачи. Если свободной задана какая-либо одна переменная $x_i(t_k)$, а остальные переменные закреплены, получим условие трансверсальности $\psi_i(t_k) = 0$.

Для нахождения условий трансверсальности также применяется прием, в котором используется вектор градиента к конечному многообразию $P(\mathbf{x})$. При этом градиент ортогонален касательной гиперплоскости L и, следовательно, в частном случае ($k = 1$) вектор $\psi(t_k)$ коллинеарен в конечной точке градиенту $\nabla h(\mathbf{x}) = \text{grad } h(\mathbf{x})$, т.е. $\psi(t_k) = \alpha \nabla h(\mathbf{x})$, где α — константа, после исключения которой найдем условия трансверсальности. В общем случае ($k > 1$) вектор $\psi(t_k)$ коллинеарен в конечной точке линейной комбинации градиентов $\nabla h_i(\mathbf{x}) = \text{grad } h_i(\mathbf{x})$, $i =$

$\overline{1, k}$, т.е. $\psi(t_k) = \sum_{i=1}^k \alpha_i \nabla h_i(\mathbf{x})$. После исключения коэффициентов α_i с учетом условия (18.15) из последнего соотношения получим условия трансверсальности.

18.1.5. Задачи Больца и Майера

Принцип максимума рассматривался выше для интегрального критерия (задача Лагранжа). Пусть задан функционал в форме Больца

$$J = \int_0^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)), \quad (18.19)$$

в котором имеется неинтегральная (терминальная) составляющая, характеризующая «стоимость» $\mathbf{x}(t)$ в конечный момент времени t_k . Минимизация критерия (18.19) выполняется при фиксированном t_k и свободном правом конце $\mathbf{x}(t_k)$.

Применение принципа максимума станет возможным, если преобразовать второе слагаемое критерия (18.19) к интегральному виду. Для этого запишем производную по t функции $F(\mathbf{x}(t))$ с учетом уравнения объекта (18.1)

$$dF/dt = (\partial F / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u}),$$

следовательно,

$$F(\mathbf{x}(t_k)) = \int_0^{t_k} (dF/dt) dt + F(\mathbf{x}_0).$$

Функционал (18.19) с учетом этого выражения примет вид

$$J = \int_0^{t_k} [f_0(\mathbf{x}, \mathbf{u}) dt + (\partial F / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u})] dt,$$

(опуская слагаемое $F(\mathbf{x}_0) = \text{const}$, не влияющее на решение).

Применительно к объекту (18.1) и этому функционалу составим гамильтониан

$$\hat{H} = -f_0(\mathbf{x}, \mathbf{u}) - (\partial F / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u}) + \hat{\psi}^T \mathbf{f}(\mathbf{x}, \mathbf{u}).$$

Запишем функцию $\tilde{H}$ для критерия (18.19) при $F(\mathbf{x}(t_k)) = 0$

$$\tilde{H} = -f_0(\mathbf{x}, \mathbf{u}) + \psi^T \mathbf{f}(\mathbf{x}, \mathbf{u}).$$

Опуская детальное обоснование, связь между векторами $\psi(t)$ и $\hat{\psi}(t)$, получим формальным путем, приравнявая функции $\tilde{H}$ и $\hat{H}$:

$$\psi = \hat{\psi} - \partial F / \partial \mathbf{x}.$$

Так как вектор $\mathbf{x}(t_k)$ свободен, согласно (18.18) $\hat{\psi}(t_k) = 0$, следовательно,

$$\psi(t_k) = -\partial F(\mathbf{x}(t_k)) / \partial \mathbf{x} = -\text{grad} F(\mathbf{x}(t_k)). \quad (18.20)$$

Таким образом, при оптимизации критерия (18.19) можно использовать функцию $\tilde{H}$, такую же как в задаче Лагранжа при $F(\mathbf{x}(t_k)) = 0$. Влияние терминальной составляющей в критерии (18.19) учитывается условием трансверсальности (18.20).

В случае применения неинтегрального критерия (задача Майера)

$$J = F(\mathbf{x}(t_k)) \rightarrow \min \quad (18.21)$$

справедливо выражение для функции $\tilde{H}$, положив в ней $f_0(\cdot) = 0$, а также введя условие трансверсальности (18.20).

Так как в задачах Майера и Больца известно конечное условие $\psi(t_k)$ в виде (18.20), они широко применяются в различных численных методах расчета оптимального управления.

18.2. ЧИСЛЕННЫЕ МЕТОДЫ РАСЧЕТА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ НА ОСНОВЕ ПРИНЦИПА МАКСИМУМА

Как отмечалось, аналитическое решение задачи оптимизации возможно лишь для достаточно простых случаев. В связи с появлением современных ЭВМ с большим быстродействием и памятью все большее применение находят численные методы оптимизации динамических объектов.

18.2.1. Решение краевой задачи путем поиска начального значения сопряженного вектора

Рассмотрим решение системы (18.13) для задачи с фиксированными граничными условиями $\mathbf{x}(0) = \mathbf{x}_0$, $\mathbf{x}(t_k) = \mathbf{x}_k$ и конечным временем t_k . Для того чтобы свести решение (18.13) к задаче Коши необходимо задать начальные условия для сопряженных переменных $\psi_i(0)$, $i = \overline{1, n}$. Выберем начальное приближение $\psi(0) = \psi^0$ и решим систему (18.13) на интервале $[0, t_k]$ при известном управлении $\mathbf{u}(\mathbf{x}, \psi)$. При $t = t_k$ находится вектор $\mathbf{x}^0(t_k)$ и вычисляется ошибка (невязка) достижения заданной конечной точки $\mathbf{e}(t_k) = \mathbf{x}^0(t_k) - \mathbf{x}_k$. Это означает, что невязка $\mathbf{e}(t_k)$ есть функция начальных значений ψ_i^0 , $i = \overline{1, n}$. Таким образом, решение задачи оптимизации свелось к поиску таких значений ψ_i^0 , $i = \overline{1, n}$, которые обращают невязки $\mathbf{e}_i(t_k)$, $i = \overline{1, n}$ в нули. Для определения нулей функции применяется метод Ньютона и его модификации. Отметим, что данный метод имеет большую чувствительность к изменению ψ_i^0 , $i = \overline{1, n}$ [51].

18.2.2. Приращение функционала

Рассмотрим итерационный процесс поиска оптимального управления $\mathbf{u}(t)$. Пусть на i -й итерации определена управляющая функция $\mathbf{u}^i(t)$. Дадим приращение управления $\Delta \mathbf{u}^i(t)$ и найдем соответствующую вариацию критерия в задаче Майера (18.21), считая, что вектор $\mathbf{u}(t)$ неограничен

$$\Delta J^i(t_k) = [\partial J / \partial \mathbf{x}^i(t_k)]^T \Delta \mathbf{x}^i(t_k) = [\partial F / \partial \mathbf{x}^i(t_k)]^T \Delta \mathbf{x}^i(t_k)$$

$$\text{или } \Delta J^i(t_k) = [\Delta \mathbf{x}^i(t_k)]^T \partial F / \partial \mathbf{x}^i(t_k). \quad (18.22)$$

Для определения приращения $\Delta \mathbf{x}^i(t)$ запишем линеаризованное дифференциальное уравнение объекта (18.1) с начальным условием $\mathbf{x}(0) = \mathbf{x}_0$

$$\dot{\mathbf{y}}^i = \mathbf{A}^i \mathbf{y}^i + \mathbf{B}^i \mathbf{v}^i, \quad (18.23)$$

где

$$\mathbf{y}^i = \Delta \mathbf{x}^i, \mathbf{v}^i = \Delta \mathbf{u}^i, \mathbf{A}^i = \partial \mathbf{f} / \partial \mathbf{x}^i, \mathbf{B}^i = \partial \mathbf{f} / \partial \mathbf{u}^i, \mathbf{y}(0) = \mathbf{0}.$$

Запишем решение уравнения (18.23), используя матрицу перехода $\Phi^i(t, \tau) = e^{\mathbf{A}^i(t-\tau)}$

$$\mathbf{y}^i(t) = \int_0^t \Phi^i(t, \tau) \mathbf{B}^i \mathbf{v}^i(\tau) d\tau. \quad (18.24)$$

Сопряженная система для уравнения (18.23) в соответствии с (18.6а) или (18.8) имеет вид

$$\dot{\psi}^i = -\mathbf{A}^{iT} \psi^i. \quad (18.25)$$

Для вектора $\psi(t)$ известно конечное состояние согласно условию трансверсальности (18.20)

$$\psi^i(t_k) = -\partial F / \partial \mathbf{x}^i(t_k). \quad (18.26)$$

Поэтому решение системы (18.25) находится в обратном времени с условием (18.26)

$$\psi^i(t) = [\Phi^i(t_k, t)]^T \psi^i(t_k). \quad (18.27)$$

Решение (18.24) при $t = t_k$ и (18.26) подставим в (18.22)

$$\Delta J^i(t_k) = - \int_0^{t_k} \Delta \mathbf{u}^{iT} \mathbf{B}^{iT} \Phi^{iT}(t_k, t) \psi^i(t_k) dt,$$

откуда с учетом решения (18.27) получим

$$\Delta J^i(t_k) = - \int_0^{t_k} \Delta \mathbf{u}^{iT} \mathbf{B}^{iT} \psi^i(t) dt. \quad (18.28)$$

Для задачи Майера и системы (18.23) определим производную

$$\partial \tilde{H} / \partial \mathbf{u}^i = [\partial f / \partial \mathbf{u}^i]^T \psi^i(t) = \mathbf{B}^{iT} \psi^i(t),$$

с учетом которой (18.28) примет вид

$$\Delta J^i(t_k) = - \int_0^{t_k} \left[\frac{\partial \tilde{H}}{\partial \mathbf{u}^i} \right]^T \Delta \mathbf{u}^i dt. \quad (18.29)$$

Выражение (18.29) дает *приращение* $\Delta J(t_k)$ *критерия* J , вызванное вариацией $\Delta \mathbf{u}^i(t)$ управления $\mathbf{u}^i(t)$, оно имеет важное значение в различных численных методах оптимизации.

18.2.3. Метод градиента для решения задачи Майера

Допустим, что управление $\mathbf{u}^i(t)$ на i -й итерации не является оптимальным и, следовательно, необходимо провести следующую итерацию, выбрав новое управление $\mathbf{u}^{i+1}(t) = \mathbf{u}^i(t) + \Delta \mathbf{u}^i(t)$. Вариация $\Delta \mathbf{u}^i(t)$ на $(i+1)$ -й итерации должна быть определена таким образом, чтобы приращение критерия $\Delta J^i(t_k)$ (18.29) было отрицательным и максимальным по величине с учетом дополнительно вводимого ограничения на изменение $\Delta \mathbf{u}^i(t)$

$$\int_0^{t_k} \Delta \mathbf{u}^{iT}(t) \mathbf{G} \Delta \mathbf{u}^i(t) dt = \varepsilon^2, \quad (18.30)$$

где $\mathbf{G}$ — постоянная матрица весовых коэффициентов, например $\mathbf{G} = \mathbf{I}$; ε^2 — заданная константа.

Решение задачи минимизации (18.29) при ограничении (18.30) дает

$$\Delta u^i(t) = -k^i G^{-1} \frac{\partial \tilde{H}}{\partial u^i}, \quad (18.31)$$

где $k^i > 0$ — постоянные коэффициенты, определяющие длину шага на $(i + 1)$ -й итерации. Таким образом, в *методе градиента* вариация управления $\Delta u^i(t)$ на каждой итерации должна выбираться *пропорциональной вектору градиента* $\partial \tilde{H} / \partial u^i$ функции Гамильтона. Трудоемкость метода значительно увеличивает необходимость перехода на каждой итерации к линеаризованной модели объекта. В связи с этим на практике часто применяют метод градиента, решая непосредственно уравнение объекта (18.1). Приведем содержание алгоритма.

1. Задается начальное приближение для управляющей функции $u^0(t)$.
2. Выполняется решение дифференциального уравнения объекта (18.1) для $u(t) = u^0(t)$ с начальным условием $x(0) = x_0$ на интервале времени $t \in [0, t_k]$.
3. При $t = t_k$ вычисляется $\psi(t_k)$ по (18.26) и решается в обратном времени от $t = t_k$ до $t = 0$ сопряженная система (18.6).
4. Используя $\psi^0(t)$, $x^0(t)$, $u^0(t)$, определяется градиент $\partial \tilde{H} / \partial u^0$.
5. В результате выбора длины шага k^0 вычисляется приращение $\Delta u^0(t)$ и новое управление $u^1(t) = u^0(t) + \Delta u^0(t)$.

Далее процедуру расчета повторяют начиная, с п.2, с заменой $u^0(t)$ на $u^1(t)$ и так далее до тех пор, пока для приращения критерия $\Delta J^i(t_k)$ не будет выполнено условие $|\Delta J^i(t_k)| \leq \delta$, где $\delta > 0$ — заданное достаточно малое число.

Приведенный алгоритм может быть применен для задачи Больца (18.19), учитывая, что в формулах (18.29), (18.30) необходимо использовать функцию Гамильтона (18.7).

Во многих реальных задачах заданы ограничения на управления, например в виде $|u_i| \leq \bar{u}_i$, $i = \overline{1, m}$. Один из часто применяемых приемов при решении таких задач состоит в добавлении к исходному функционалу специально подобранных штрафных функций аналогично тому, как это выполняется в нелинейном программировании. Аналогичный подход используют в случае фиксированного конечного условия $x(t_k) = x_k$.

Отдельного рассмотрения требуют задачи, в которых конечное время t_k не задано и, следовательно, оно должно быть найдено в процессе оптимизации. Прямой метод решения может состоять из последовательности решений оптимизационных задач с различными фиксированными значениями t_k в соответствии с градиентным алгоритмом и определением решения с минимальным t_k .

Отметим, что метод градиента имеет быструю сходимость в начальной стадии расчета, которая, однако, существенно снижается в окрестности оптимального решения. Более эффективные алгоритмы численного расчета оптимального управления основаны на методе сопряженных градиентов, а также на использовании второй вариации критерия [87].

18.2.4. Метод последовательных приближений (МПП)

Рассмотрим задачу Майера. В методе последовательных приближений (МПП) [89] первые три шага вычислений на каждой итерации совпадают с градиентным методом (п.18.2.3.). Отличие от метода градиента заключается в том, что управление $u^i(t)$ на i -й итерации определяется непосредственно из условия максимума функции Гамильтона

$$\tilde{H}^i = \psi^{i^*} f(x^i, u^i)$$

с учетом ограничения $u^i(t) \in U$. Обозначим это управление как $u^{i^*}(t)$. В простейшей форме алгоритма МПП на $(i+1)$ -й итерации принимается

$$u^{i+1}(t) = u^{i^*}(t), \quad i = 0, 1, 2, \dots$$

В виду возможной расходимости расчета предложены различные модификации простейшего алгоритма [89]. Рассмотрим две модификации.

1. Управление на $(i+1)$ -й итерации формируется в виде линейной выпуклой комбинации нового управления $u^{i^*}(t)$ и старого $u^i(t)$:

$$u^{i+1}(t) = \alpha u^{i^*}(t) + (1 - \alpha) u^i(t), \quad \alpha \in [0, 1].$$

Параметр α на каждой итерации подбирается из условия минимума критерия J по α . При этом требуется пересчитывать только систему уравнений объекта (18.1); интегрирование сопряженной системы и вычисление максимума функции Гамильтона выполняются только один раз на каждой итерации. Отметим, что для данной модификации приближение к релейному управлению (характерному для задачи оптимального быстрогодействия) не будет идти в классе релейных управлений.

2. Управление на $(i+1)$ -й итерации формируется путем сочетания нового $u^{i^*}(t)$ и старого $u^i(t)$ управлений таким образом, что каждое из управлений действует на «своем» отрезке времени в соответствии с алгоритмом:

$$u^{i+1}(t) = \begin{cases} u^{i^*}(t), & t \in [t', t''] \\ u^i(t), & t \notin [t', t''] \end{cases}$$

$$[t', t''] \subset [0, t_k], i = 0, 1, 2, \dots$$

Подбор $[t', t'']$ осуществляется так, чтобы обеспечить строгое убывание функционала $J^{i+1} < J'$. Возможен более простой вариант: положив $t'' = t_k$, варьировать только одним параметром $t' \in [0, t_k]$. В этой модификации МПП на $(i + 1)$ -й итерации сохранится релейный характер управления, если он имел место хотя бы для одного из $u^{*}(t)$, $u'(t)$.

18.2.5. Преобразование к задаче нелинейного программирования

Наряду с применением численных методов на базе вариационных методов имеется направление, ориентированное на преобразование исходной непрерывной задачи оптимизации к задаче математического программирования, для решения которой разработано большое число пакетов прикладных программ. Рассмотрим такое преобразование для объекта (18.1) с ограничением (18.2) и критерия (18.3) при заданном конечном времени t_k и фиксированных граничных условий $x(0) = x_0$, $x(t_k) = x_k$.

Введем шаг дискретизации по времени $h = t_k / N$, где N — число точек решения на интервале $[0, t_k]$. Заменим интеграл в критерии (18.3) суммой

$$J = h \sum_{i=0}^{N-1} f_0(x_i, u_i). \quad (18.32)$$

Дифференциальное уравнение (18.1) преобразуется к разностному уравнению

$$x_{i+1} = x_i + hf(x_i, u_i), \quad (18.33)$$

которое получается в результате применения простейшей разностной схемы

$$\dot{x} = (x_{i+1} - x_i) / h = f(x_i, u_i).$$

Выразим решение уравнения (18.33) x_i через $x(0) = x_0$ и $u_0, \dots, u_{i-1}$

$$\begin{aligned} x_1 &= x_0 + hf(x_0, u_0) = F_1(u_0), \\ x_2 &= F_1(u_0) + hf[F_1(u_0), u_1] = F_2(u_0, u_1), \\ &\dots \end{aligned} \quad (18.34)$$

$$\begin{aligned} x_N &= F_{N-1}(u_0, \dots, u_{N-2}) + hf[F_{N-1}(u_0, \dots, u_{N-2}), u_{N-1}] = \\ &= F_N(u_0, \dots, u_{N-1}), \end{aligned} \quad (18.35)$$

где функции $F_i(\cdot)$ находятся в процессе конкретных вычислений для заданных x_0 , h и $f(x, u)$. Так как $x_N = x_k$ то решение (18.35) представляет

собой ограничение типа равенства. После подстановки решений $x_1, \dots, x_{N-1}$ (18.34) в (18.32) критерий становится функцией только векторов $u_0, \dots, u_{N-1}$

$$J = h\{f_0(u_0) + f_0[F_1(u_0), u_1] + \dots + f_0[F_{N-1}(u_0, \dots, u_{N-2}), u_{N-1}]\}$$

или после преобразования

$$J = \Phi(u_0, u_1, \dots, u_{N-1}). \quad (18.36)$$

Таким образом, осуществлен переход к задаче нелинейного программирования: требуется определить последовательность управлений $\{u_0, u_1, \dots, u_{N-1}\}$, удовлетворяющих ограничениям $u_i \in U$, $i = \overline{0, N-1}$, ограничению равенству (18.35) и доставляющих минимум целевой функции (18.36). Недостаток метода состоит в значительном усложнении решения задачи с увеличением размерности исходной системы уравнений и интервала $[0, t_k]$.

18.3. МЕТОД ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Наряду с принципом максимума широкое применение для решения задач оптимального управления находит метод динамического программирования, разработанный Р. Беллманом. В отличие от принципа максимума этот метод оптимизации приводит к управлению как функции вектора состояния $u = u(t, x(t))$, т.е. в виде обратных связей, решая тем самым задачу синтеза оптимального управления. Метод полностью обоснован для дискретных систем и в меньшей степени для непрерывных систем.

18.3.1 Принцип оптимальности

Вывод условий оптимальности динамического программирования базируется на принципе оптимальности, который, в свою очередь, использует свойства оптимальной траектории $x^*(t)$, $u^*(t)$, $t \in [0, t_k]$. Это свойство состоит в том, что любой участок оптимальной траектории, например от t' до t_k , также является оптимальной траекторией, где $0 < t' < t_k$. Также важно свойство решения дифференциального уравнения (18.1): решение $x(t)$ при $t > t'$, $0 < t' \leq t_k$, а также значение критерия (18.3) определяются начальным состоянием $x(t')$ и управлением $u(t)$, $t \geq t'$. Приведем теперь принцип оптимальности: каково бы ни было первоначальное состояние и управление на начальном этапе $t \in [0, t']$, последующее управление и состояние при $t > t'$ должны быть оптимальными относительно состояния, в которое приходит система в результате начального этапа.

Рассмотрим применение этого достаточно общего принципа для отдельных классов задач оптимизации.

18.3.2. Дискретные детерминированные системы

Пусть объект управления описывается векторным разностным уравнением

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n, \mathbf{u}_n), \quad (18.37)$$

которое может непосредственно отражать реальную работу объекта, а также может быть получено в результате перехода от непрерывного описания к дискретному.

Качество работы системы оценивается критерием

$$J = T \sum_{n=0}^{N-1} f_0(\mathbf{x}_{n+1}, \mathbf{u}_n) + F(\mathbf{x}_N) \rightarrow \min, \quad (18.38)$$

получаемым путем перехода от непрерывного критерия Больца (18.19).

При формировании критерия учтена особенность работы дискретного объекта, а именно в соответствии с уравнением (18.37) управление $\mathbf{u}_n$ в момент времени nT , $n = \overline{0, N-1}$, определяет значение состояния $\mathbf{x}_{n+1}$ в момент $(n+1)T$ при известном $\mathbf{x}_n$. Поэтому в выражение (18.38) не входит $\mathbf{x}_0$, как не зависящее от $\mathbf{u}_n$.

Критерий (18.38) представляет собой функцию многих переменных $\mathbf{x}_n$, $\mathbf{u}_n$. Пусть заданы начальные условия $\mathbf{x}(0) = \mathbf{x}_0$, $t_k = NT$, а конечное условие $\mathbf{x}_N$ — свободно. Требуется определить последовательность управлений $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{N-1}$ и соответствующую ей последовательность состояний $\mathbf{x}_1, \dots, \mathbf{x}_N$, доставляющую минимум функционалу (18.38). Управление может быть ограниченным $\mathbf{u}_n \in U$. Для простоты рассмотрим объект управления первого порядка с одной управляющей координатой.

В методе динамического программирования минимизация функции многих переменных (18.38) сводится к многошаговой оптимизации, так что на каждом шаге необходимо решать более простую задачу определения минимума функции одной переменной.

Введем функцию Беллмана, принимая $T = 1$

$$S_{N-j}(\mathbf{x}_{N-j}) = \min_{\mathbf{u}_{N-1}, \dots, \mathbf{u}_{N-j}} \left[\sum_{i=N-j}^{N-1} f_0(\mathbf{x}_{i+1}, \mathbf{u}_i) + F(\mathbf{x}_N) \right],$$

представляющую собой минимальное значение функционала на j последних шагах $(N-j)T \leq t \leq NT$.

Начнем решение задачи с последнего шага, т.е. $t \in [(N-1)T, NT]$. Начальному моменту времени $(N-1)T$ для этого шага соответствует состояние x_{N-1} , которое может принимать различные значения из возможной области изменения x . Для задачи Больца конечное значение x_N свободно, поэтому, также как для x_{N-1} , необходимо при выборе решения на последнем шаге учесть все допустимые значения x_N . Обычно рассматривают множество дискретных значений x , заданных с определенным шагом. Значение критерия (18.38) на последнем шаге определяется значениями x_{N-1} , x_N и управлением на этом шаге:

$$J_N = f_0(x_N, u_{N-1}) + F(x_N), \quad (18.39)$$

где управление u_{N-1} находится из уравнения объекта

$$x_N = f(x_{N-1}, u_{N-1}). \quad (18.40)$$

Управление u_{N-1} следует выбирать таким образом, чтобы обеспечить минимум (18.39) при выполнении условия (18.40). Отметим, что для задачи Лагранжа с фиксированным значением x_N при $F(x_N) = 0$ в критерии (18.38) исключается перебор значений x_N и, следовательно, не требуется минимизации (18.39), так как для каждого значения x_{N-1} управление u_{N-1} находится из уравнения (18.40). Обозначим управление, определенное на последнем шаге, через $u_{N-1}^*(x_{N-1})$; ему соответствует значение функционала (18.39):

$$\min_{u_{N-1}} J(x_{N-1}, u_{N-1}) = J(x_{N-1}, u_{N-1}^*(x_{N-1})) = S_{N-1}(x_{N-1}). \quad (18.41)$$

Значения критерия $S_{N-1}(x_{N-1})$ и управления $u_{N-1}^*(x_{N-1})$ необходимо занести в память ЭВМ.

Рассмотрим теперь момент времени $t_{N-2} = (N-2)T$ с начальным значением состояния x_{N-2} и найдем минимум функционала для двух последних шагов $(N-1)$ -го и N -го:

$$S_{N-2}(x_{N-2}) = \min_{u_{N-1}, u_{N-2}} [f_0(x_{N-1}, u_{N-2}) + f_0(x_N, u_{N-1}) + F(x_N)].$$

Согласно принципу оптимальности запишем

$$S_{N-2}(x_{N-2}) = \min_{u_{N-2}} [f_0(x_{N-1}, u_{N-2}) + S_{N-1}(x_{N-1})], \quad (18.42)$$

где $S_{N-1}(x_{N-1})$ было вычислено согласно (18.41). Управление u_{N-2} должно минимизировать все выражение под знаком минимума, так как необходимо учесть, что значение x_{N-1} зависит от управления u_{N-2} на предпоследнем шаге в соответствии с решением (18.37)

$$x_{N-1} = f(x_{N-2}, u_{N-2}). \quad (18.42a)$$

В результате минимизации (18.42) с учетом (18.42а) найдем $u_{N-2}^*(x_{N-2})$ — оптимальное управление на отрезке времени $[(N-2)T, (N-1)T]$, доставляющее вместе с управлением $u_{N-1}^*(x_{N-1})$ минимум функционалу на двух последних шагах. Значения $S_{N-2}(x_{N-2})$ и $u_{N-2}^*(x_{N-2})$ необходимо запомнить.

Продолжая аналогичным образом расчеты на других шагах, получим рекуррентное соотношение Беллмана для момента времени $(N-j)T$

$$S_{N-j}(x_{N-j}) = \min_{u_{N-j}} \left[f_0(x_{N-j+1}, u_{N-j}) + S_{N-j+1}(x_{N-j+1}) \right], \quad (18.43)$$

где $x_{N-j+1} = f(x_{N-j}, u_{N-j})$, $j = \overline{1, N}$.

Путем минимизации выражения в квадратных скобках определим оптимальное управление $u_{N-j}^*(x_{N-j})$ на отрезке времени $[(N-j)T, (N-j+1)T]$. Вычисляя последовательно согласно уравнению (18.43) $S_{N-j}(x_{N-j})$ для $j = \overline{1, N}$, приходим к управлению $u_0^*(x_0)$ на первом шаге, причем одновременно с этим получим и $S_0(x_0)$ — минимальное значение критерия (18.38) при оптимальном управлении на всем заданном интервале времени $[0, NT]$. Просмотрев результаты расчетов от $t = 0$ с начальным условием $x(0) = x_0$ к конечному моменту времени NT , определим оптимальное управление как функцию времени $u^*(n)$, $n = 0, 1, \dots, N-1$, а также как функцию состояния $u^*(x_n)$.

Несмотря на наглядность метода динамического программирования, его применение для численного расчета требует большого объема памяти ЭВМ, резко возрастающего с увеличением порядка системы уравнений объекта, а также размерности вектора управления u . Если объект имеет m управляющих координат, то на каждом шаге оптимизации необходимо определять экстремум функции m переменных относительно $u_1, \dots, u_m$. Для снижения количества просматриваемых в процессе расчета точек разработаны различные приемы, такие как полиномиальная аппроксимация решения, изменение периода дискретизации по времени [26].

18.3.3. Непрерывные детерминированные системы

Получим условия оптимальности в непрерывной форме для объекта (18.1) и критерия (18.19). Считаем заданными начальное состояние $x(0) = x_0$ и время перехода t_k , конечное состояние $x(t_k)$ — свободно.

Введем функцию Беллмана

$$S(\tau, x(\tau)) = \min_{u(\tau)} \left[\int_{\tau}^{t_k} f_0(x, u) dt + F(x(t_k)) \right], \quad (18.44)$$

которая представляет минимальное значение функционала (18.19) на отрезке $[\tau, t_k]$, определяемое начальным состоянием $x(\tau)$, $0 \leq \tau \leq t_k$. Разо-

бьем отрезок $[\tau, t_k]$ на два: $[\tau, \tau + \Delta]$ и $[\tau + \Delta, t_k]$, $\Delta > 0$. В выражении (18.44) интеграл представим в виде суммы двух интегралов

$$S(\tau, \mathbf{x}(\tau)) = \min_{\mathbf{u}(t)} \left[\int_{\tau}^{\tau+\Delta} f_0(\mathbf{x}, \mathbf{u}) dt + \int_{\tau+\Delta}^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)) \right]. \quad (18.45)$$

Поведение управления $\mathbf{u}(t)$ на отрезке $[\tau + \Delta, t_k]$ не влияет на величину первого интеграла в (18.45), поэтому на этом отрезке $\mathbf{u}(t)$ выбирается так, чтобы обеспечить

$$\min_{\mathbf{u}(t)} \left[\int_{\tau+\Delta}^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)) \right].$$

Учитывая обозначения (18.44), перепишем выражение (18.45)

$$S(\tau, \mathbf{x}(\tau)) = \min_{\mathbf{u}(t)} \left[\int_{\tau}^{\tau+\Delta} f_0(\mathbf{x}, \mathbf{u}) dt + S(\tau + \Delta, \mathbf{x}(\tau + \Delta)) \right], \quad t \in [\tau, \tau + \Delta], \quad (18.46)$$

Управление $\mathbf{u}(t)$, $t \in [\tau, \tau + \Delta]$ необходимо выбирать таким образом, чтобы минимизировать все выражение в квадратных скобках на отрезке $[\tau, t_k]$, так как значение $\mathbf{x}(\tau + \Delta)$, от которого зависит $S(\tau + \Delta, \mathbf{x}(\tau + \Delta))$, определяется управлением $\mathbf{u}(t)$, $t \in [\tau, \tau + \Delta]$ согласно решению уравнения (18.1).

Предположим, что $S(\tau + \Delta, \mathbf{x}(\tau + \Delta))$ — непрерывная дифференцируемая функция. Представим ее в виде ряда при $\Delta \rightarrow 0$, учитывая только члены первого порядка малости

$$S(\tau + \Delta, \mathbf{x}(\tau + \Delta)) \cong S(\tau, \mathbf{x}(\tau)) + \partial S / \partial t \cdot \Delta + (\partial S / \partial \mathbf{x})^T \Delta \mathbf{x}. \quad (18.47)$$

Из уравнения объекта (18.1) приближенно получим

$$\Delta \mathbf{x} \approx \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau)) \Delta. \quad (18.48)$$

Подставим (18.47) в выражение (18.46) с учетом (18.48)

$$S(\tau, \mathbf{x}(\tau)) \cong \min_{\mathbf{u}(t)} \left[f_0(\mathbf{x}, \mathbf{u}) \Delta + S(\tau, \mathbf{x}(\tau)) + \partial S / \partial t \cdot \Delta + (\partial S / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u}) \Delta \right], \quad (18.49)$$

где было принято, что $\int_{\tau}^{\tau+\Delta} f_0(\mathbf{x}, \mathbf{u}) dt \approx f_0(\mathbf{x}(\tau), \mathbf{u}(\tau)) \Delta$.

В равенстве (18.49) функции $S(\tau, \mathbf{x}(\tau))$ и $\partial S / \partial t$ не зависят от $\mathbf{u}(t)$, поэтому их можно вынести из под знака минимума, что приводит к функциональному уравнению Беллмана после сокращения на Δ и замены τ на t :

$$\partial S / \partial t = - \min_{\mathbf{u}(t)} \left[f_0(\mathbf{x}, \mathbf{u}) + (\partial S / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u}) \right], \quad (18.50)$$

которое дает необходимое условие экстремума функционала (18.19).
Функция S должна удовлетворять граничному условию

$$S(t_k, \mathbf{x}(t_k)) = F(\mathbf{x}(t_k)). \quad (18.51)$$

Для определения оптимального управления в *открытой* области изменения $\mathbf{u}(t)$ вычислим производную по $\mathbf{u}$ выражения под знаком минимума и приравняем ее нулю

$$\partial f_0 / \partial \mathbf{u} + (\partial \mathbf{f} / \partial \mathbf{u})^B \cdot \partial S / \partial \mathbf{x} = 0,$$

откуда получим

$$\mathbf{u} = \mathbf{u}(\mathbf{x}, \partial S / \partial \mathbf{x}). \quad (18.52)$$

После подстановки управления (18.52) в уравнение Беллмана (18.50) приходим к дифференциальному уравнению в частных производных $\Phi(\mathbf{x}, \partial S / \partial \mathbf{x}, \partial S / \partial t) = 0$ относительно функции $S(t, \mathbf{x}(t))$ с граничным условием (18.51). Знание функции $S(t, \mathbf{x}(t))$ позволяет определить оптимальное управление согласно (18.52) непосредственно как функцию вектора состояния $\mathbf{x}$. Аналитически найти решение уравнения в частных производных во многих случаях невозможно и поэтому применяют различные численные методы для нахождения $S(t, \mathbf{x}(t))$ и управления (18.52).

Рассмотрим ситуацию, когда конечное время t_k *свободно*. В этом случае для объекта (18.1) и критерия (18.19) при $F(\mathbf{x}(t_k)) = 0$ функция Беллмана не зависит явно от времени t , т. е. $\partial S / \partial t = 0$, и уравнение (18.50) принимает вид

$$0 = \min_{\mathbf{u}(t)} \left[f_0(\mathbf{x}, \mathbf{u}) + (\partial S / \partial \mathbf{x})^T \mathbf{f}(\mathbf{x}, \mathbf{u}) \right]$$

с граничным условием $S(\mathbf{x}(t_k)) = 0$. Сравнение этого уравнения с условием оптимальности принципа максимума (18.12) и функцией Гамильтона (18.7) приводит к соотношениям между вспомогательными переменными принципа максимума ψ_i и частными производными $\partial S / \partial x_i$, $i = \overline{1, n}$ метода динамического программирования

$$\psi_i = -\partial S / \partial x_i, \quad i = \overline{1, n},$$

$$\psi_0 = -1.$$

Отметим, что при выводе уравнения Беллмана существенным было предположение о дифференцируемости функции $S(t, \mathbf{x}(t))$, которое не выполняется для многих практических задач, например для задач, имеющих управление в виде кусочно-постоянной функции времени. Указанное ограничение на функцию $S(t, \mathbf{x}(t))$ может быть ослаблено, а именно, $S(t, \mathbf{x}(t))$ может иметь разрывы частных производных на некотором множестве точек [17].

Пусть объект описывается уравнением

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) + \mathbf{G}\xi(t), \quad (18.53)$$

где $\xi(t)$ — возмущение в виде случайного процесса типа белого шума с характеристиками

$$M[\xi(t)] = 0, M[\xi(t)\xi^T(\tau)] = \mathbf{V}\delta(t - \tau),$$

где M — символ математического ожидания, $\delta(t - \tau)$ — дельта-функция, $\mathbf{V}$ — матрица интенсивности белого шума.

Для каждой отдельно взятой реализации процесса $\mathbf{x}(t)$, $\mathbf{u}(t)$ критерий (18.19) принимает случайное значение. Поэтому обычно выполняют усреднение получаемых результатов и рассматривают стохастический критерий

$$J = M \left[\int_0^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)) \right]. \quad (18.54)$$

Допустим, что управление формируется как функция текущего состояния $\mathbf{x}(t)$, т.е. $\mathbf{u} = \Phi(t, \mathbf{x}(t))$. Как и в детерминированном случае (п.18.3.3), запишем функцию Беллмана

$$S(\tau, \mathbf{x}(\tau)) = \min_{\mathbf{u}(t)} M \left\{ \left[\int_{\tau}^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)) \right] / \mathbf{x} \right\}, t \in [\tau, t_k], \quad (18.55)$$

где $M[(\cdot)/\mathbf{x}]$ означает условное математическое ожидание от функции $(\cdot)$ при $\mathbf{x}(t) = \mathbf{x}$, $t \in [\tau, t_k]$.

Введем момент времени $\tau + \Delta \leq t_k$, $\Delta > 0$, тогда

$$S(\tau, \mathbf{x}(\tau)) = \min_{\mathbf{u}(t)} M \left\{ \left[\int_{\tau}^{\tau+\Delta} f_0(\mathbf{x}, \mathbf{u}) dt + \int_{\tau+\Delta}^{t_k} f_0(\mathbf{x}, \mathbf{u}) dt + F(\mathbf{x}(t_k)) \right] / \mathbf{x} \right\}.$$

На основе принципа оптимальности, опуская преобразования, учитывая свойства условного математического ожидания, и имея в виду выражение (18.46), получим при достаточно малом Δ [55, 73]

$$S(\tau, \mathbf{x}(\tau)) = \min_{\mathbf{u}(t)} \{ f_0(\mathbf{x}, \mathbf{u})\Delta + M[S(\tau + \Delta, \mathbf{x}(\tau + \Delta))/\mathbf{x}] \}, t \in [\tau, \tau + \Delta], \quad (18.56)$$

Функцию $M[S(\tau + \Delta, \mathbf{x}(\tau + \Delta))/\mathbf{x}]$ представим в виде ряда, используя стохастическое дифференцирование и пренебрегая членами высокого порядка малости

$$M[S(\tau + \Delta, \mathbf{x}(\tau + \Delta))/\mathbf{x}] \cong S(\tau, \mathbf{x}(\tau)) + \partial S / \partial t \cdot \Delta + (\partial S / \partial \mathbf{x})^T \Delta \mathbf{x} + \frac{1}{2} \text{tr} \left(\mathbf{G} \mathbf{V} \mathbf{G}^T \frac{\partial^2 S}{\partial \mathbf{x}^2} \right),$$

где $\text{tr}(\cdot)$ — след матрицы, т.е. сумма элементов ее главной диагонали.

После подстановки последнего выражения в уравнение (18.56), выполнив те же действия, что и при переходе от (18.49) к (18.50), приходим к стохастическому уравнению Беллмана

$$\partial S / \partial t = - \min_{u(t)} \left[f_0(x, u) + (\partial S / \partial x)^T f(x, u) + \frac{1}{2} \text{tr} \left(G V G^T \frac{\partial^2 S}{\partial x^2} \right) \right] \quad (18.57)$$

с граничным условием $S(t_k, x(t_k)) = F(x(t_k))$. Схема нахождения оптимального управления из уравнения (18.57) соответствует детерминированному случаю (п.18.3.3).

Глава 19

СИСТЕМЫ, ОПТИМАЛЬНЫЕ ПО БЫСТРОДЕЙСТВИЮ

19.1. ОПРЕДЕЛЕНИЕ АЛГОРИТМА УПРАВЛЕНИЯ

Рассмотрим задачу управления линейным стационарным объектом

$$\dot{x} = Ax + Bu, \quad (19.1)$$

где A, B — матрицы размером $n \times n, n \times m$ соответственно; x — n -мерный вектор состояния; u — m -мерный вектор управления.

Заданы ограничения на изменения управляющих переменных

$$|u_i| \leq \bar{u}_i, \quad i = \overline{1, m}, \quad (19.2)$$

которые в m -мерном пространстве образуют параллелепипед.

Качество работы оценивается функционалом

$$J = \int_0^{t_k} 1 \cdot dt = t_k. \quad (19.3)$$

Необходимо найти $u(t)$, обеспечивающее минимальное время перехода (19.3) объекта (19.1) из начального состояния $x(0) = x_0$ в конечное $x(t_k) = x_k$. По смыслу задачи конечный момент времени t_k следует считать свободным.

Для нахождения алгоритма управления применим принцип максимума. Запишем гамильтониан (18.7), учитывая, что $\psi_0 = -1$, (п.18.1.2)

$$\tilde{H} = -1 + \psi^T Ax + \psi^T Bu, \quad (19.4)$$

причем $\max \tilde{H} = 0$ на оптимальной траектории для любого $t \in [0, t_k]$ (п.18.1.2). Для задачи быстрогодействия целесообразно применять более простую функцию

$$H = \psi^T A x + \psi^T B u, \quad (19.5)$$

тогда $H = \tilde{H} + 1$ и

$$\max H = 1 \quad (19.6)$$

для любого $t \in [0, t_k]$ на оптимальной траектории. Соотношение (19.6) есть условие *транскверсальности* в задаче быстрогодействия. Оно совместно с граничными условиями для вектора x образует полный набор условий, необходимых для решения двухточечной граничной задачи.

Запишем уравнение для вспомогательной переменной $\psi(t)$ согласно (18.8)

$$\dot{\psi} = -\partial H / \partial x = -A^T \psi. \quad (19.7)$$

Уравнение (19.7) называется *сопряженным* по отношению к уравнению объекта (19.1). Решение уравнения (19.7) может быть записано через матрицу перехода (см. п. 2.7.2)

$$\psi(t) = e^{-A^T t} \psi(0). \quad (19.8)$$

Сопряженная система имеет следующие свойства.

1. Пусть α_i — собственные числа матрицы A , т.е. корни характеристического уравнения объекта (19.1) $\det(sI - A) = 0$. Тогда собственные числа λ_i матрицы $-A^T$, т.е. корни характеристического уравнения $\det(sI + A^T) = 0$ сопряженной системы (19.7) связаны с корнями α_i соотношением $\lambda_i = -\alpha_i$, $i = \overline{1, n}$.

2. Вспомогательные переменные $\psi_i(t) \neq 0$, $i = \overline{1, n}$ одновременно для любого $t \in [0, t_k]$, что проверяется анализом функции H (19.5) и условия (19.6).

3. Переменные $\psi_i(t)$ как решения (19.8) системы (19.7) есть непрерывные функции времени.

При решении задачи в общем случае векторного управления $m > 1$ представим матрицу B в виде совокупности векторов-столбцов $B = [b_1, \dots, b_m]$. После этого уравнение объекта и функция H примут вид

$$\dot{x} = Ax + b_1 u_1 + \dots + b_m u_m, \quad (19.9)$$

$$H = \psi^T Ax + \psi^T b_1 u_1 + \dots + \psi^T b_m u_m. \quad (19.10)$$

Введем функции

$$\sigma_i(t) = b_i^T \psi(t) = \psi^T(t) b_i, \quad i = \overline{1, m}. \quad (19.11)$$

Из выражения (19.10) выпишем слагаемые, явно зависящие от u

$$H_1 = \sigma_1(t) u_1 + \dots + \sigma_m(t) u_m, \quad (19.12)$$

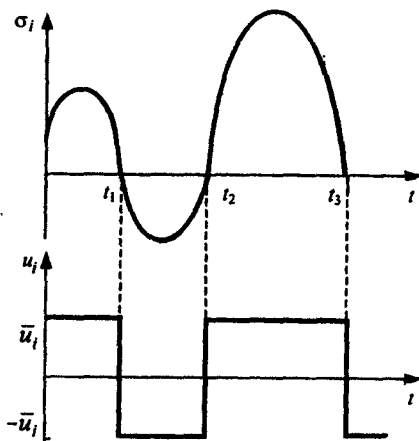


Рис. 19.1. Формирование оптимального по быстродействию управления

причем максимум функций H и H_1 по u_i достигается одновременно. Таким образом, требуется решать задачу

$$H_1(\sigma_i, u_i) \rightarrow \max_{|u_i| \leq \bar{u}_i}.$$

Из структуры функции H_1 следует, что ее максимум достигается, если каждое из слагаемых в выражении (19.12) принимает максимальное значение, т.е.

$$\sigma_i(t)u_i \rightarrow \max, |u_i| \leq \bar{u}_i, i = \overline{1, m}. \quad (19.13)$$

Функция в левой части (19.13) линейно зависит от управления u_i , поэтому максимум ее достигается на границе допустимой области (19.2). Учитывая это, запишем условие оптимальности управления в зависимости от знака $\sigma_i(t)$:

$$u_i = \bar{u}_i \text{sign } \sigma_i(t), i = \overline{1, m} \quad (19.14)$$

при условии, что функция $\sigma_i(t) = 0$ в отдельные моменты времени, называемые точками (моментами) переключения управления $u_i(t)$ с $\bar{u}_i$ на $-\bar{u}_i$ или с $-\bar{u}_i$ на $\bar{u}_i$, $i = \overline{1, m}$. Таким образом, оптимальное по быстродействию управление $u_i(t)$, $i = \overline{1, m}$ имеет *разрывный* характер (рис. 19.1). Задача оптимального быстродействия, для которой справедлив алгоритм управления (19.14), называется *нормальной задачей*.

19.1.1. Вырожденность управления

Возможен случай, когда на каком-либо ненулевом отрезке времени $[t', t''] \subset [0, t_1]$ функция

$$\sigma_i(t) = 0, \quad i = \overline{1, m}, \quad (19.15)$$

Тогда функция $\sigma_i u_i = 0$ при любом $-u \leq u \leq u$, $i = \overline{1, m}$, т.е. управление не определено. Такое управление называют *вырожденным*.

Для определения условий возникновения вырожденности управления запишем систему линейных алгебраических однородных уравнений относительно $\psi(t)$, беря в качестве исходного уравнение (19.15):

$$\begin{cases} \sigma_i(t) = \mathbf{b}_i^T \boldsymbol{\psi} = 0, \\ \dot{\sigma}_i(t) = \mathbf{b}_i^T \dot{\boldsymbol{\psi}} = -\mathbf{b}_i^T \mathbf{A}^T \boldsymbol{\psi} = 0, \\ \dots\dots\dots, \\ \sigma_i^{(n-1)}(t) = (-1)^{n-1} \mathbf{b}_i^T \mathbf{A}^{T^{n-1}} \boldsymbol{\psi} = 0. \end{cases} \quad (19.16)$$

Так как вектор $\psi(t) \neq 0$ для любого $t \in [0, t_k]$, при решении системы (19.16) важно найти нетривиальное решение $\psi(t) \neq 0$. Такое решение существует, если

$$\det G_i = \det [\mathbf{b}_i \quad \mathbf{A}\mathbf{b}_i \dots \mathbf{A}^{n-1}\mathbf{b}_i] = 0, \quad i = \overline{1, m}.$$

Следовательно, вырожденность управления $u(i)$ возникает только в том случае, когда хотя бы по одной управляющей координате объект (19.1) не полностью управляем. Вырожденность отсутствует, и, следовательно, задача является *нормальной* с алгоритмом управления (19.14), если $\det G_i \neq 0$ для всех $i = \overline{1, m}$, что соответствует условию нормальности объекта (§ 3.13). Нормальность объекта есть необходимое и достаточное условие отсутствия вырожденности задачи оптимального быстродействия.

19.1.2. Теорема об n интервалах

Условие оптимальности (19.14) определяет качественный характер алгоритма управления. Более точное представление об особенностях алгоритма дает *теорема об n интервалах*, которая формулируется следующим образом.

Пусть система (19.1) нормальна и собственные числа $\alpha_1, \dots, \alpha_n$ матрицы A — действительные числа. Тогда максимальное число интервалов знакопостоянства управления не превосходит n для всех $i = \overline{1, m}$, соответственно число переключений не более $(n - 1)$. На каждом ин-

тервале знакопостоянства управления $u_i(t)$ принимает максимальное по величине значение $u_i, i = \overline{1, m}$.

Дадим пояснения к этой теореме, опуская полное доказательство [17, 32]. Из условия (19.11) следует, что нули функции $\sigma_i(t)$ зависят от решения $\psi(t)$, которое в покоординатной форме запишем в следующем виде (при отсутствии кратных корней):

$$\psi_i(t) = \sum_{j=1}^n c_{ij} e^{-\alpha_j t}, \quad i = \overline{1, n}.$$

Тогда

$$\sigma_i(t) = \mathbf{b}_i^T \psi(t) = \sum_{j=1}^n c_{ij} e^{-\alpha_j t}, \quad i = \overline{1, m}, \quad (19.17)$$

где c_{ij}, c_j — постоянные, определяемые в процессе конкретных вычислений, при заданных $\lambda_j = -\alpha_j, j = \overline{1, n}$ они зависят от начального условия $\psi(0)$ в (19.8). Поскольку $\psi(0)$, а следовательно, и c_j неизвестны, теорема устанавливает не конкретное число интервалов для заданных граничных условий, а только *верхнюю границу* числа интервалов. Анализ показывает [17, 32], что функция $\sigma_i(t)$ (19.17) для любого $i = \overline{1, m}$, как сумма n экспонент, имеет число нулей, т.е. число переключений управления $u_i(t)$, не более $(n - 1)$.

Эта теорема имеет важное практическое значение, позволяя разработать методы расчета моментов переключения управляющих воздействий, а также синтезировать управление с обратными связями при построении замкнутых систем, оптимальных по быстродействию.

Если имеется хотя бы одна пара комплексно-сопряженных корней $\alpha_{1,2} = \alpha \pm j\beta$, теорема в общем случае неверна, так как при этом в функцию $\sigma_i(t)$ войдут слагаемые вида $ce^{-\alpha t} \sin(\beta t + \varphi)$. Число нулей $\sigma_i(t)$ зависит от граничных условий и может быть как меньше или равно n , так и больше n , оставаясь конечным.

Выделим практически важную ситуацию, когда граничные условия $\mathbf{x}(0), \mathbf{x}(t_r)$ соответствуют установившемуся решению уравнения объекта (19.1) (задача позиционного управления). Анализ фазовых траекторий для системы второго порядка показывает, что оптимальное по времени управление может быть реализовано в соответствии с теоремой об n интервалах как двухинтервальное.

Пример 1. Рассмотрим слабо демпфированный объект, описываемый уравнениями

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -\omega_0^2 x_1 - 2\xi\omega_0 x_2 + ku, \quad |u| \leq 1,$$

где $\omega_0 = 2 \text{ с}^{-1}, \xi = 0,1, k = 8$.

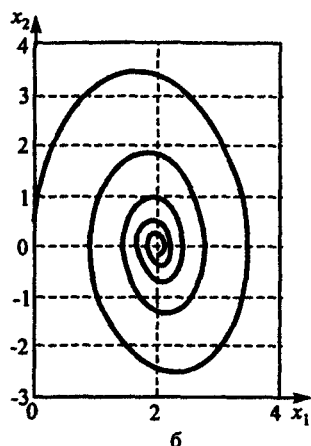
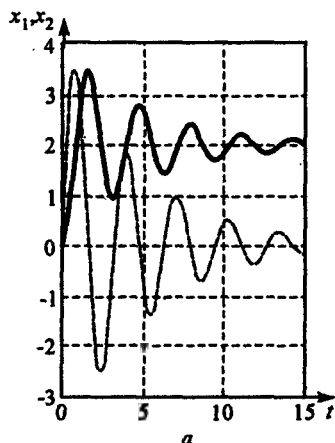


Рис. 19.2. Переходные процессы и фазовая траектория колебательного объекта

Переходные процессы $x_i(t)$ являются слабозатухающими колебательными (рис.19.2, а). Фазовая траектория объекта при $u = 1$ с начальным условием $x(0) = 0$ имеет вид спирали (рис.19.2, б). Найдем управление с минимальным временем t_k для граничных условий $x_1(0) = x_2(0) = 0$, $x_1(t_k) = 1,8$, $x_2(t_k) = 0$. Расчет моментов переключения t_1 и окончания процесса t_k , основанный на решении уравнения объекта в прямом и обратном времени, будет приведен в п.19.2.3. Значения моментов времени $t_1 = 0,608$, $t_k = 0,836$, координаты точки переключения $x_1(t_1) = 1,21$, $x_2(t_1) = 3,31$. Переходные процессы $x_1(t)$, $x_2(t)$ (рис.19.3, а) и фазовая траектория (рис.19.3, б), состоящая из двух ветвей при $u = 1$ и $u = -1$, показывают высокую эффективность оптимального по быстродействию алгоритма управления, а именно, значительное уменьшение длительности переходного процесса при отсутствии перерегулирования.

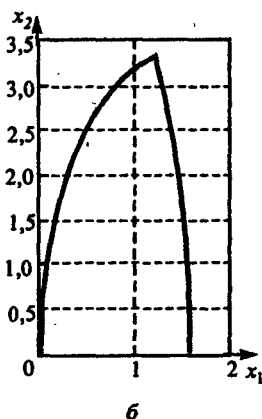
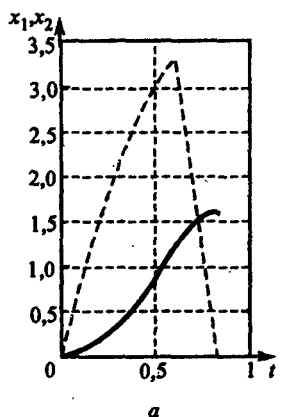


Рис. 19.3. Оптимальные по быстродействию переходные процессы и фазовая траектория для колебательного объекта

19.2. РАСЧЕТ МОМЕНТОВ ПЕРЕКЛЮЧЕНИЯ

Знание максимального числа интервалов управления для объекта с действительными собственными числами матрицы A в соответствии с теоремой об n интервалах, дает возможность рассчитать моменты переключения $t_1, \dots, t_{n-1}$ и конечное время $t_k = t_n$. Рассмотрим два подхода, обусловленные видом математической модели объекта управления.

19.2.1. Описание объекта в пространстве состояний

Пусть $m = 1$ (скалярное управление)

$$\dot{x} = Ax + bu, \quad |u| \leq \bar{u}. \quad (19.18)$$

Запишем решение уравнения (19.18) при $t = t_n$ для релейного управления с моментами переключения $t_1, \dots, t_{n-1}$ предполагая, что на первом интервале $u = \bar{u}$

$$\begin{aligned} x(t_n) = & e^{At_n} x_0 + \bar{u} \int_0^{t_1} e^{A(t_n-\tau)} b d\tau - \bar{u} \int_{t_1}^{t_2} e^{A(t_n-\tau)} b d\tau + \dots \\ & \dots + (-1)^{n-1} \bar{u} \int_{t_{n-1}}^{t_n} e^{A(t_n-\tau)} b d\tau = x_k, \end{aligned} \quad (19.19)$$

где $t_n = t_k$ — конечное время переходного процесса; $t_1, \dots, t_{n-1}$ — моменты переключений; x_0, x_k — граничные условия. Так как x — n -мерный вектор, то равенству (19.19) соответствуют n нелинейных алгебраических (трансцендентных) уравнений относительно $t_1, \dots, t_{n-1}, t_n$. Решение должно быть физически реализуемым, поэтому необходимо учесть условия $0 \leq t_i \leq t_{i+1}$, $i = \overline{1, n-1}$. Кроме того, в тех случаях, когда затруднительно выбрать знак управления на первом интервале, необходимо выполнить решение (19.19) для двух управляющих последовательностей, отличающихся чередованием знаков. В общем случае решение (19.19) возможно только численным методом на ЭВМ.

19.2.2. Описание объекта в виде уравнения вход-выход

Объект управления описывается одним уравнением n -го порядка с действительными корнями его характеристического полинома

$$a_n x^{(n)} + \dots + a_{n-1} \dot{x} + a_0 x = bu, \quad |u| \leq \bar{u}. \quad (19.20)$$

От уравнения вход-выход (19.20) перейдем к эквивалентной системе уравнений в пространстве состояний (19.18). Выбрав, например, в качестве переменных состояния $x, \dot{x}, \dots, x^{(n-1)}$, получим систему в каноническом управляемом базисе (п.2.12.1). Затем для нахождения $t_p, i = \overline{1, n}$ используется система (19.19).

Другой подход базируется на свойстве непрерывности решения $x(t)$ уравнения (19.20), а также производных $\dot{x}(t), \dots, x^{(n-1)}(t)$ для любого $t \in [0, t_k]$, включая моменты переключения $t_p, i = \overline{1, n}$. Рассмотрим метод стыковки решений для вычисления $t_p, i = \overline{1, n}$. Возьмем момент переключения $t_q, q = \overline{1, n}$, предполагая, что на q -м интервале $u = \bar{u}$. Тогда для него можно записать решение (при $b = a_0$)

$$\begin{aligned} x(t_q) &= \sum_{i=1}^n c_i^q e^{a_i t_q} + \bar{u}, \\ x^{(j)}(t_q) &= \sum_{i=1}^n c_i^q a_i^j e^{a_i t_q}, \quad j = \overline{1, n-1}, \end{aligned} \quad (19.21)$$

где верхний индекс для c_i^q обозначает номер интервала управления. Одновременно момент времени t_q есть начало $(q+1)$ -го интервала, на котором $u(t) = -\bar{u}$. Этому моменту соответствуют решения

$$\begin{aligned} x(t_q) &= \sum_{i=1}^n c_i^{q+1} e^{a_i t_q} - \bar{u}, \\ x^{(j)}(t_q) &= \sum_{i=1}^n c_i^{q+1} a_i^j e^{a_i t_q}, \quad j = \overline{1, n-1}. \end{aligned} \quad (19.22)$$

Учитывая непрерывность переменных при $t = t_q$, решения (19.21) и (19.22) следует покоординатно приравнять, после чего получим систему

$$\begin{aligned} \sum_{i=1}^n (c_i^{q+1} - c_i^q) e^{a_i t_q} &= 2\bar{u}, \\ \sum_{i=1}^n (c_i^{q+1} - c_i^q) a_i^j e^{a_i t_q} &= 0, \quad j = \overline{1, n-1}, \end{aligned} \quad (19.23)$$

Системы такого вида формируются для каждого $q = \overline{1, n-1}$. Кроме того, необходимо найти решения для $t = 0$

$$\begin{aligned} \sum_{i=1}^n c_i^1 &= x_0 - \bar{u}, \\ \sum_{i=1}^n c_i^1 a_i^j &= x_0^{(j)}, \quad j = \overline{1, n-1}, \end{aligned} \quad (19.24)$$

а также для $t = t_n = t_k$

$$\sum_{i=1}^n c_i^n e^{a_i t_n} + \bar{u} = x_k, \quad (19.25)$$

$$\sum_{i=1}^n c_i^n a_i^j e^{a_i t_n} = x_k^{(j)}, \quad j = \overline{1, n-1},$$

предполагая, что на первом и последнем интервалах $u = \bar{u}$. Постоянные c_i^q , $q = \overline{1, n}$ последовательно исключаются из решений, начиная с уравнений (19.24), (19.25), в результате чего придем к системе нелинейных алгебраических (трансцендентных) уравнений

$$\eta_i(t_1, \dots, t_n, x_0, x_k, a_1, \dots, a_n, \bar{u}) = 0, \quad i = \overline{1, n}, \quad (19.26)$$

из которой находятся моменты $t_1, \dots, t_n$. Решение системы (19.26) также как и системы (19.19) в общем случае производится численными методами на ЭВМ. В [54] имеется таблица, где даны уравнения вида (19.26) для ряда конкретных описаний объекта управления ($n = 2+4$).

Пример 2. Рассмотрим объекты, характеристическое уравнение которых имеет нулевые корни. В таком случае возможно аналитическое решение. Пусть дифференциальное уравнение объекта второго порядка имеет вид

$$T\ddot{x} + \dot{x} = ku, \quad |u| \leq \bar{u},$$

корни его характеристического уравнения равны $\alpha_1 = -1/T = \alpha$, $\alpha_2 = 0$. Запишем уравнения (19.26) для данного случая

$$(1 + \dot{x}_k/b)e^{-\alpha t_2} - 2e^{-\alpha t_1} + (1 - \dot{x}_0/b) = 0,$$

$$t_2 - 2t_1 + (x_k - x_0)/b - (\dot{x}_k - \dot{x}_0)/\alpha b = 0.$$

Решение этих уравнений сводится к решению квадратного уравнения относительно переменной $z = e^{\alpha t}$, выразив t_2 через t_1 из второго уравнения и подставив в первое.

Другой объект представляет собой последовательное включение двух интегрирующих звеньев и описывается дифференциальным уравнением

$$\ddot{x} = ku, \quad |u| \leq \bar{u}$$

с двумя нулевыми корнями характеристического уравнения $\alpha_1 = \alpha_2 = 0$. Моменты времени t_1, t_2 находятся из уравнений

$$t_2 - 2t_1 + (\dot{x}_k - \dot{x}_0)/b = 0, \quad t_2^2 - 2t_1^2 + 2\dot{x}_k/b - 2(x_k - x_0)/b = 0.$$

Для обоих объектов приняты обозначения: $b = k\bar{u}$, $x_0, \dot{x}_0$ — начальные, а $x_k, \dot{x}_k$ — конечные условия; знак первого интервала принят положительным. В случае позиционного управления в приведенных выше уравнениях следует положить $\dot{x}_0 = \dot{x}_k = 0$.

В общем случае аналитическое решение получено для объекта, структурно представляющего собой последовательное соединение n интеграторов, с дифференциальным уравнением

$$x^{(n)} = ku, \quad |u| \leq \bar{u}.$$

Для позиционного управления моменты переключения находятся из выражения

$$t_i = 4\sqrt[n]{(n-1)!|(x_k - x_0)/4b|} \cdot \sin^2 \frac{i\pi}{2n}, \quad i = \overline{1, n}.$$

19.2.3. Численный метод расчета на ЭВМ оптимальных по быстродействию процессов

Ввиду трудности решения уравнений (19.19), (19.26) разрабатываются методы численного расчета моментов t_i , $i = \overline{1, n}$ на ЭВМ. Для линейных систем имеются методы, ориентированные на поиск начального условия для сопряженного вектора $\psi(0)$, входящего в функции $\sigma_i(t)$, $i = \overline{1, n}$ (19.11) (метод Нойштадта) [87]. Другие методы [40] осуществляют непосредственно варьирование моментов переключений. Рассмотрим достаточно простой и наглядный метод расчета моментов переключений и оптимальных процессов для линейных и нелинейных объектов второго порядка, не требующий задания начального приближения и проведения какого-либо итерационного процесса [31].

Пусть объект описывается уравнениями

$$\dot{x}_1 = f_1(x_1, x_2, u), \quad \dot{x}_2 = f_2(x_1, x_2, u), \quad (19.27)$$

где x_1 — выходная координата, $|u| \leq u$. Считаем, что объект полностью управляем и оптимальность достигается при релейном управлении с числом интервалов не более двух. Для определения моментов t_1, t_2 используем прямое движение из начальной точки $x(0) = x_0$ и обратное движение из конечной $x(t_k) = x_k$. Для определенности примем, что на первом интервале $u = u$, а на втором $u = -u$.

Для объекта (19.27) запишем две системы дифференциальных уравнений

$$dx_1/dt = f_1(x, u), \quad dx_2/dt = f_2(x, u), \quad (19.28)$$

$$dz_1/d\tau = -f_1(z, u), \quad dz_2/d\tau = -f_2(z, u), \quad (19.29)$$

где переменные x_1, x_2 соответствуют движению на первом интервале управления в прямом времени t при $u = u$, $z^T = [z_1, z_2]$ изменению координат объекта на втором интервале при $u = -u$ в обратном времени $\tau = t_k - t$. Решение (19.28) выполняется из начальной точки $x(0) = x_0$, а (19.29) из конечной точки $z_1(0) = x_{1k}$, $z_2(0) = x_{2k}$.

Пусть для простоты $x_2(0) = x_{2k}$. Используя один из численных методов решения дифференциальных уравнений (метод Эйлера–Коши, Рунге–Кутты и др.) выполняется одновременное решение систем (19.28), (19.29) соответственно из заданных начальной и конечной точек. В про-

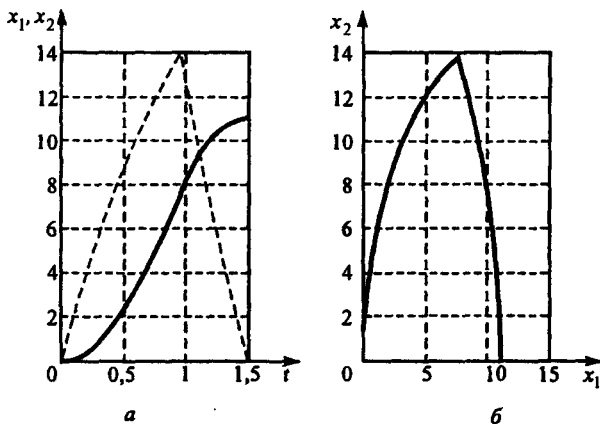


Рис. 19.4. Результаты численного расчета оптимальных по быстродействию процессов для нелинейного объекта

цессе решения на каждом шаге интегрирования производится сравнение координат x_1 и z_1 траекторий прямого и обратного движений. При этом необходимо учесть разные скорости протекания процессов на первом и втором интервалах управления. При выполнении условия $|x_{1i} - z_{1i}| \leq \varepsilon$, где ε — требуемая точность достижения точки переключения, решение задачи заканчивается. Варианты расчета при различных видах функций $f_i(x, u)$ и $x_2(0) \neq x_{2k}$, а также развитие данного метода на нелинейные объекты более высокого порядка приведены в [31].

Пример 3. Объект задан уравнениями

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -0,5x_1 - 0,05x_1x_2 + u, \quad |u| \leq 20.$$

Граничные условия: $x(0) = 0$, $x_{1k} = 11$, $x_{2k} = 0$. Шаг интегрирования равен 0,01, метод Эйлера–Коши. Результаты вычислений: $t_1 = 0,958$, $t_k = 1,489$ при $\varepsilon = 0,001$. На рис. 19.4, а, б приведены оптимальные переходные процессы и фазовые траектории соответственно.

Информация о моментах переключения $t_1, \dots, t_n$ дает возможность реализовать программное во времени управление $u(t)$, используя для этого аппаратные или программные средства. Такой способ построения оптимально по времени системы применяется редко в связи с известными недостатками, присущими разомкнутым системам управления. Более совершенными характеристиками обладают замкнутые системы с информацией о состоянии объекта управления.

19.3. ЗАМКНУТЫЕ ОПТИМАЛЬНЫЕ ПО БЫСТРОДЕЙСТВИЮ СИСТЕМЫ

При построении оптимальных по времени замкнутых систем управления используется информация о переменных состояния объекта $x_i(t)$, $i = \overline{1, n}$. Для реализации требуемого алгоритма управления применяют *функцию переключения*

$$\mu = \mu(x_k, x, \bar{u}, a),$$

где a — параметры, входящие в уравнения объекта. Так как $x(t)$ — функция времени, $\mu = \mu(x_k, x(t), \bar{u}, a)$ есть неявная функция времени. Сигнал, соответствующий $\mu(x(t))$, подается на вход идеального релейного элемента (РЭ) (рис. 19.5).

Вид функции переключения должен быть определен таким образом, чтобы нули функции $\mu[x(t)]$ совпали с нулями функции (19.11) $\sigma(t) = b^T \psi(t)$ (для скалярного управления, $m = 1$). При этом не требуется точного совпадения функций $\mu(t)$ и $\sigma(t)$ на интервалах времени между их нулями.

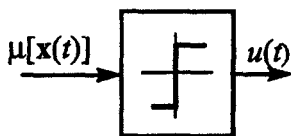


Рис. 19.5. Формирование релейного управления из функции переключения

Для синтеза функции переключения обычно рассматривают движение объекта в пространстве состояний $x_i(t)$, $i = \overline{1, n}$. Определение функции $\mu(x)$ в аналитическом виде для объектов порядка $n > 2$ связано в общем случае с большими трудностями, так как невозможно найти уравнения траекторий в пространстве состояний. Один из приемов для объектов с действительными попарно различными собственными числами α_i , $i = \overline{1, n}$ матрицы A состоит в преобразовании ее к диагональной $\Lambda = \text{diag}[\alpha_1, \dots, \alpha_n]$. Уравнения объекта при таком преобразовании примут вид

$$\dot{y}_i(t) = \alpha_i y_i + d_i u, \quad i = \overline{1, n}.$$

Опуская вычисления, запишем уравнения проекций траектории в n -мерном пространстве состояний на координатные плоскости (y_i, y_n) , $i = \overline{1, n-1}$ при $u = \text{const}$

$$y_i = (y_{i0} + c_i u) \left(\frac{y_n + c_n u}{y_{n0} + c_n u} \right)^{\gamma_i} - c_i u, \quad i = \overline{1, n-1}, \quad (19.30)$$

где $c_i = d_i / \alpha_i$; $c_n = d_n / \alpha_n$; $\gamma_i = \alpha_i / \alpha_n$; y_{i0} , y_{n0} — координаты точки, через которую проходит проекция (19.30). Уравнения проекций (19.30) по-

звolyют синтезировать функцию переключения для объектов высокого порядка [57]. Отметим также применение ЭВМ для определения функции $\mu(x)$, используя приближенные методы расчета траекторий [32,40].

19.4. СИНТЕЗ ЗАМКНУТОЙ ОПТИМАЛЬНОЙ ПО БЫСТРОДЕЙСТВИЮ СИСТЕМЫ ВТОРОГО ПОРЯДКА

Рассмотрим объект управления, описываемый уравнением

$$T_0 \ddot{x} + \dot{x} = ku, \quad |u| \leq \bar{u}. \quad (19.31)$$

Примерами объектов с таким уравнением могут быть привод станка, печь с интегрирующим исполнительным механизмом и др.

Требуется за минимальное время t_k перевести объект (19.31) из начального состояния $x(0) = x_0$ на траекторию входного сигнала $g(t) = g_0 + g_1 t$. Сигнал $g(t)$ можно считать программой, которую должен воспроизвести выход объекта (19.31). В частном случае при $g_1 = 0$ получим ступенчатое воздействие величиной g_0 . Потребуем, чтобы при $t > t_k$ осуществлялось точное воспроизведение сигнала задания, т.е. $x(t) \equiv g(t)$.

Введем новые переменные: ошибку воспроизведения $e = g - x$ и производную $\dot{e} = \dot{g} - \dot{x}$. Тогда $\ddot{e} = \ddot{g} - \ddot{x} = -\ddot{x}$ и уравнение объекта (19.31) относительно ошибки $e(t)$ примет вид

$$T_0 \ddot{e} + \dot{e} = g_1 - ku, \quad |u| \leq \bar{u}. \quad (19.32)$$

В момент времени $t = t_k$ переменные $e(t_k) = \dot{e}(t_k) = 0$. Таким образом, необходимо осуществить переход за минимальное время из начального состояния $e(0) = e_0 = g_0 - x(0)$, $\dot{e}(0) = \dot{e}_0 = g_1 - \dot{x}(0)$ в начало координат $e(t_k) = \dot{e}(t_k) = 0$.

От уравнения (19.32) перейдем к эквивалентной системе уравнений

$$\dot{e}_1 = e_2, \quad \dot{e}_2 = -e_2/T_0 - ku/T_0 + g_1/T_0, \quad (19.33)$$

где $e_1 = e$, $e_2 = \dot{e}$.

Назовем воздействие $g(t) = g_0 + g_1 t$ допустимым, если выполняется условие $|g_1| < k\bar{u}$, иначе при $t > t_k$ объект не может отслеживать сигнал $g(t)$. По теореме об n интервалах оптимальное по быстродействию управление $u(t)$ системой (19.33) имеет не более двух интервалов знакопостоянства и одного переключения знака.

Найдем дифференциальное уравнение фазовых траекторий, разделив первое уравнение системы (19.33) на второе,

$$\frac{de_1}{de_2} = -T_0 \frac{e_2}{e_2 + \sigma k\bar{u} - g_1}, \quad \sigma = \pm 1,$$

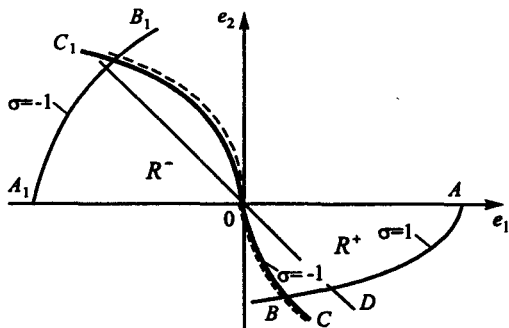


Рис. 19.6. Фазовые траектории оптимальной по времени системы второго порядка

выполнив интегрирование которого, получим уравнение для семейства фазовых траекторий

$$e_1 = -T_0 e_2 + T_0 (\sigma k \bar{u} - g_1) \ln |e_2 + \sigma k \bar{u} - g_1| + D, \quad (19.34)$$

где постоянная D определяется условием прохождения фазовой траектории через заданную точку.

Построим фазовую траекторию AB из начальной точки A для первого интервала управления $\sigma = 1$ (рис. 19.6).

Затем находим фазовую траекторию $C0$, проходящую на втором интервале $\sigma = -1$ через начало координат. В этом случае

$$D = -T_0 (\sigma k \bar{u} - g_1) \ln |\sigma k \bar{u} - g_1|$$

и уравнение (19.34) примет вид

$$e_1 = -T_0 e_2 + T_0 (\sigma k \bar{u} - g_1) \ln |(e_2 + \sigma k \bar{u} - g_1) / (\sigma k \bar{u} - g_1)|. \quad (19.35)$$

Аналогичным образом могут быть получены траектории $A_1 B_1$, $C_1 0$ для другой последовательности управления $\{-1, 1\}$. Траектории $C0$ и $C_1 0$, проходящие через конечную точку 0 , образуют линию переключения $C0C_1$ — геометрическое место точек, соответствующих моментам переключения t_1 для различных $e_1(0)$, $e_2(0)$.

Найдем уравнение линии переключения $C0C_1$. Для ее ветви $C0$ $\sigma = -1$, для ветви $C_1 0$ $\sigma = 1$. Система управления должна автоматически выбирать требуемый знак $u(t)$ или σ . Правильный выбор ветви $C0$ или $C_1 0$ может быть сделан, если принять $\sigma = \text{sign } e_2$. После подстановки $\sigma = \text{sign } e_2$ в (19.35) получим уравнение линии переключения

$$e_1 = -T_0 y + T_0 (v \text{sign } y - g_1) \ln |(y + v \text{sign } y - g_1) / (v \text{sign } y - g_1)|, \quad (19.36)$$

где $y = e_2$, $v = k \bar{u}$.

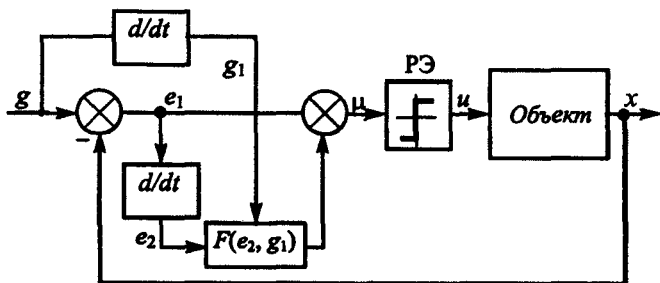


Рис. 19.7. Структурная схема оптимальной по времени системы второго порядка

Линия переключения COC_1 делит всю плоскость e_1, e_2 на две области: R^+ , где $\sigma = 1$, и R^- , где $\sigma = -1$. В процессе движения из начальной точки A изображающая точка пересекает линию COC_1 и в этот момент происходит смена знака управления. Уравнение (19.36) дает возможность найти *функцию переключения*

$$\mu = e + T_0 y - T_0 (v \operatorname{sign} y - g_1) \ln |(y + v \operatorname{sign} y - g_1) / (v \operatorname{sign} y - g_1)| \quad (19.37)$$

или в сокращенной форме

$$\mu = e + F(y, g_1, v),$$

где $F(\cdot)$ — сумма второго и третьего слагаемых в (19.37).

Из способа образования функции μ следует, что $\mu = 0$ на линии переключения COC_1 . Анализ показывает, что $\mu > 0$ в области R^+ и $\mu < 0$ в области R^- . Следовательно, функция μ (19.37) обеспечит требуемый алгоритм оптимального управления и может быть применена в качестве функции переключения. Структурная схема оптимальной системы приведена на рис. 19.7.

Особенность ее состоит в том, что в ней наряду с каналом ошибки $e_1 = g - x$ имеется прямая связь по производной $\dot{g}(t)$, т.е. оптимальная система является *комбинированной*. Эта схема обеспечивает оптимальные по времени процессы для линейного $g(t) = g_1 t$, ступенчатого $g(t) = g_0 1(t)$ сигналов задания и их линейной комбинации. Если убрать связь по $\dot{g}(t)$, тогда сохранится оптимальность только для ступенчатых сигналов задания.

Основная трудность при построении оптимальной системы связана с реализацией нелинейной функции $F(y, g_1, v)$, которая может быть выполнена на основе современных аналоговых микросхем, а также программным путем при включении ЭВМ в контур управления.

Отметим особенность работы системы с функцией μ (19.37). Применение этой функции переключения дает так называемую *техничес-*

кую реализацию оптимального алгоритма. Это обусловлено тем, что точное движение по линии $СОС_1$ невозможно, так как при этом $\mu = 0$. Реально движение происходит за линией переключения на достаточно малом расстоянии, в зависимости от зоны нечувствительности РЭ. Отклонение реальных параметров объекта от расчетных приводит к возникновению скользящего режима, что увеличивает время переходного процесса. Анализ фазовых траекторий, приведенных на рис. 19.6, показывает, что функция μ (19.37) необходима прежде всего для определения момента переключения, т.е. на первом интервале управления. Для организации движения на втором интервале функция переключения может быть заменена другой, при которой достигается больший сигнал на входе РЭ. В частности, целесообразно применение линейной функции $\mu = e_1 + ly$, которой соответствует прямая OD (см. рис. 19.6). Коэффициент l часто выбирают экспериментально так, чтобы достигнуть желаемого качества процесса в области скользящего режима.

19.4.1. Учет изменения параметров объекта

Пусть на вход системы действует ступенчатый сигнал задания $g(t) = g_0 1(t)$. С учетом этого запишем (19.37) в виде

$$\mu = e + T_0 y - T_0 \sigma v \ln |(y + \sigma v) / \sigma v|, \quad \sigma = \text{sign} y. \quad (19.38)$$

Введем новую переменную $\chi = |y|/v$, учитывая, что $|y| = y \text{sign} y$. Для упрощения последующего анализа алгоритмов представим приближенно третье слагаемое в (19.38) в виде ряда

$$T_0 \sigma v \ln |(1 + \chi)| \cong T_0 \sigma v \left(\chi - \frac{\chi^2}{2} + \frac{\chi^3}{3} - \dots + (-1)^{m+1} \frac{\chi^m}{m} \right). \quad (19.39)$$

Так как $\sigma v \chi = y$, то, ограничиваясь двумя членами ряда (19.39), получим из (19.38) упрощенную функцию переключения

$$\mu \cong e + T_0 y |y| / 2v. \quad (19.40)$$

Предположим, что параметры объекта могут в процессе работы изменяться *неконтролируемым* образом. Рассмотрим несколько способов адаптации в зависимости от того, какие параметры объекта (19.31) являются переменными, предполагая, что $k = k_1 k_2$ (рис. 19.8, а). Кроме того, считаем, что имеется возможность измерения $x_1, x_2, \dot{x}$. Обозначим $T = T_0$.

1. $k_1 = \text{var}, k_2 = \text{var}, T_0 = \text{const}$. Введем звено, являющееся моделью апериодического звена, входящего в состав объекта (рис. 19.8, а), где $k_1 k_2 = k$. При нулевых начальных условиях $x_2(0) = x_{\mu}(0) = 0$ получим соотношение

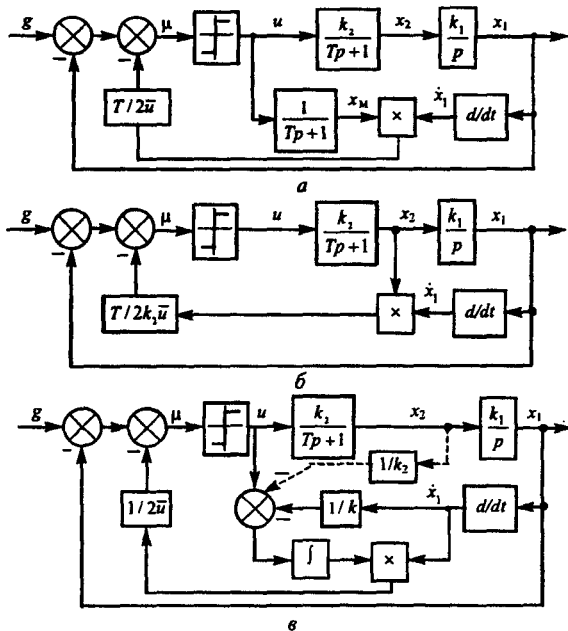


Рис. 19.8. Структурные схемы оптимальных по времени систем второго порядка при изменении параметров: а — k_1, k_2 ; б — k_1 ; в — $T = T_0$

$x_m = x_2/k_2$. Так как $x_2 = \dot{x}/k_1$, то $x_m = \dot{x}/k$. Тогда вместо (19.40) получим функцию переключения

$$\mu \equiv e - T_0 \dot{x} |x_m| / 2\bar{u},$$

имея в виду, что $y = e_2 = \dot{e} = -\dot{x}$. Эта функция переключения, а значит и алгоритм управления, не зависит от коэффициентов k_1, k_2 .

2. $k_1 = \text{var}, k_2 = \text{const}, T_0 = \text{const}$. Считаем, что можно измерить координату x_2 и вычислить (измерить) производную $\dot{x}$. В этом случае получим $x_2 = \dot{x}/k_1$ и, следовательно, после подстановки в (19.40) придем к функции переключения

$$\mu \equiv e - T_0 \dot{x} |x_2| / 2k_2\bar{u},$$

не зависящей от коэффициента k_1 , рис. 19.8, б,

3. $k_1 = \text{const}, k_2 = \text{const}, T_0 = \text{var}$. Из уравнения объекта (19.31) получим

$$T_0 \ddot{x} / k = u - \dot{x} / k.$$

После интегрирования при нулевых начальных условиях придем к соотношению

$$T_0 \dot{x} / k = \int (u - \dot{x} / k) dt,$$

что приводит к функции переключения

$$\mu = e - \dot{x}|z|/2\bar{u},$$

не зависящей от постоянной времени T_0 (рис. 19.8, в), где $z = T_0 \dot{x} / k$ находится в результате операции интегрирования.

4. $k_1 = \text{var}$, $k_2 = \text{const}$, $T_0 = \text{var}$. Аналогично п. 3 из уравнения объекта (19.31) после интегрирования определим

$$z' = T_0 \dot{x} / k = \int (u - x_2 / k_2) dt.$$

Следовательно, $\mu \equiv e - \dot{x}|z'|/2\bar{u}$ и не зависит от k_1 и T_0 (см. рис. 19.8, в — штриховая линия).

5. $k_1 = \text{var}$, $k_2 = \text{var}$, $T_0 = \text{var}$. Запишем решение для переменной $x_2(t) = k_2 \bar{u}(1 - e^{-t/T_0})$ и ее производной $\dot{x}_2(t) = k_2 \bar{u} e^{-t/T_0} / T_0$, откуда при $t = 0$ получим $\dot{x}_{20} = k_2 \bar{u} / T_0$ и $T_0 / k_2 \bar{u} = 1/\dot{x}_{20}$. После подстановки в (19.40) найдем функцию переключения $\mu \equiv e - \dot{x}|x_2|/2\dot{x}_{20}$, не зависящую от k_1 , k_2 и T_0 . Сложность данного способа адаптации вызвана необходимостью дифференцирования переменной $x_2(t)$ и запоминания начального значения $\dot{x}_{20}$ на время, равное длительности первого интервала.

Анализ показывает, что эти способы адаптации применимы и в случае более точного представления функции переключения (19.38) с большим числом членов ряда (19.39).

Отметим, что кроме рассмотренных здесь способов адаптации разрабатываются более общие методы, учитывающие алгоритмы оценивания как вектора состояния, так и вектора параметров объекта управления и формирующие на основе этой информации алгоритмы оптимального управления.

Для нелинейных объектов второго порядка функция переключения может быть получена путем решения уравнений (19.27) или (19.29) в обратном времени из конечной точки x_k . Вычисленные таким образом точки x_{1i} , x_{2i} , $i = 1, 2, \dots$ принадлежат линии переключения, приближенное уравнение которой $x_1 = \eta(x_2, x_k)$ находится путем аппроксимации таблично заданной функции заранее выбранным выражением, например полиномом соответствующей степени.

Продолжение примера 3. Для иллюстрации используем результаты численного расчета оптимальных процессов, приведенные в примере 3 (п. 19.2.3). Применяя метод наименьших квадратов при выполнении полиномиальной аппроксимации функции, заданной массивами точек $\{x_1^1, \dots, x_1^N\}$, $\{x_2^1, \dots, x_2^N\}$, принадлежащих фазовой траектории второго интервала, т.е. линии переключения, получим уравнение

$$x_1 = x_{1k} - a_1 x_2 - a_2 x_2^2 - a_3 x_2^3,$$

где $a_1 = 0,082$, $a_2 = 0,0212$, $a_3 = -0,0003$. Максимальная погрешность аппроксимации равна 0,00766 в точке $x_2 = 14$.

От линии переключения перейдем к функции переключения

$$\mu(x_{1k}, x_1, x_2) = x_{1k} - x_1 - a_1 x_2 - a_2 x_2^2 - a_3 x_2^3,$$

что позволит реализовать оптимальное по времени управление в замкнутой системе

$$u(t) = \overline{\text{sign}} \mu(x_{1k}, x_1, x_2).$$

Необходимо заметить, что алгоритмы, оптимальные по быстродействию, наиболее широко применяются при отработке достаточно больших начальных отклонений. В области малых рассогласований, когда в большей степени влияют неизбежные на практике неточности в описании объекта, а также погрешности в реализации алгоритма, часто переходят на другие законы управления, в частности на линейные.

19.5. ОПТИМАЛЬНЫЕ ПО БЫСТРОДЕЙСТВИЮ СИСТЕМЫ С ОГРАНИЧЕННЫМИ РЕСУРСАМИ

В системах, оптимальных по быстродействию, управление имеет релейный характер, принимая максимальное по величине значение на каждом интервале знакопостоянства в соответствии с (19.14). Такое управление сопровождается значительными затратами энергии (топлива) за время переходного процесса. В связи с этим важное значение имеет задача определения алгоритмов управления, которые при оптимизации учитывали бы не только время перехода из начальной точки в конечную, но и требуемые при этом расходы энергии (топлива). Особенно это актуально для объектов с автономными источниками энергии, с ограниченным запасом топлива, а также при использовании дорогостоящего сырья.

Дополнительно к функционалу быстродействия (19.3) введем критерий

$$J' = \int_0^{t_k} f_0(u) dt \rightarrow \min. \quad (19.41)$$

При оптимизации по расходу топлива применяется критерий с функцией

$$f_0(u) = \sum_{i=1}^m r_i |u_i|, \quad (19.42)$$

а по расходу энергии с положительно-определенной функцией

$$f_0(u) = u^T R u \quad (19.43)$$

или в частном случае диагональной матрицы R

$$f_0(u) = \sum_{i=1}^m r_i u_i^2, \quad (19.43a)$$

где $r_i > 0$, $i = \overline{1, m}$ — весовые коэффициенты.

Учет времени процесса и расхода энергии (топлива) приводит к двухкритериальной задаче оптимизации, в которой требуется найти управление, доставляющее минимум критериям (19.3) и (19.41). Эти критерии являются противоречащими друг другу, т.е. увеличение одного из них сопровождается уменьшением другого. Решение данной двухкритериальной задачи часто сводят к решению задачи с одним общим аддитивным критерием

$$J = \int_0^{t_k} [\lambda + f_0(u)] dt, \quad (19.44)$$

где $\lambda \geq 0$ — весовой коэффициент, посредством которого задают определенную степень компромисса между временем процесса и затратами на управление.

19.5.1. Оптимизация по критерию расхода топлива

Уравнение объекта управления представим в виде, аналогичном задаче быстрогодействия (19.9)

$$\dot{x} = Ax + b_1 u_1 + \dots + b_m u_m. \quad (19.45)$$

Для этого уравнения и критериев (19.41), (19.42) запишем функцию Гамильтона

$$\tilde{H} = -r_1 |u_1| - \dots - r_m |u_m| + \psi^T Ax + \psi^T b_1 u_1 + \dots + \psi^T b_m u_m,$$

Выделим из функции $\tilde{H}$ слагаемые, явно зависящие от u_i , $i = \overline{1, m}$,

$$\tilde{H}_i = (-r_i |u_i| + \psi^T b_i u_i) + \dots + (-r_m |u_m| + \psi^T b_m u_m), \quad (19.46)$$

причем максимумы функций $\tilde{H}$ и $\tilde{H}_i$ достигаются одновременно. При определении максимума $\tilde{H}_i$ (19.46) необходимо обеспечить максимум каждого выражения в круглых скобках:

$$-r_i |u_i| + \psi^T b_i u_i \rightarrow \max_{|u_i| \leq \bar{u}_i}, \quad i = \overline{1, m}. \quad (19.47)$$

После деления на $r_i > 0$ и введения функций

$$\sigma_i(t) = b_i^T \psi(t) / r_i, \quad i = \overline{1, m} \quad (19.48)$$

задача (19.47) преобразуется к виду

$$-|u_i| + \sigma_i(t) u_i \rightarrow \max_{|u_i| \leq \bar{u}_i}, \quad i = \overline{1, m}. \quad (19.49)$$

Определим оптимальное управление u^* , путем анализа целевой функции в задаче (19.49) в зависимости от $\sigma_i(t)$, учитывая, что $-|u_i| \leq 0, i = \overline{1, m}$

$$u_i^*(t) = \begin{cases} \bar{u}_i \operatorname{sign} \sigma_i(t), & \text{если } |\sigma_i(t)| > 1, \\ 0, & \text{если } |\sigma_i(t)| < 1, i = \overline{1, m}, \end{cases} \quad (19.50)$$

Отдельного рассмотрения требует случай, когда $\sigma_i(t) = 1, i = \overline{1, m}$. Если $\sigma_i(t) = 1, i = \overline{1, m}$ только в отдельные моменты времени на отрезке времени $[0, t_k]$, то оптимальное управление с учетом (19.50) может быть записано в краткой форме:

$$u_i(t) = \bar{u}_i \operatorname{dez} \sigma_i(t), i = \overline{1, m}, \quad (19.51)$$

где dez означает функцию, соответствующую характеристике релейного звена с зоной нечувствительности, равной единице. Такое управление называется *невыврожденным* или *нормальным*.

Другая ситуация соответствует *вырожденному* управлению, возникающему при условии, что $|\sigma_i(t)| = 1, i = \overline{1, m}$ для любого t на ненулевом отрезке времени $[t', t''] \subset [0, t_k]$. В этом случае из анализа функции цели в (19.49) следует, что при $\sigma_i(t) = 1, i = \overline{1, m}$ управление $u_i(t)$ может принимать любое значение от 0 до $\bar{u}_i$, а при $\sigma_i(t) = -1$ — любое значение от $\bar{u}_i$ до 0.

Для определения условий возникновения *вырожденности* запишем равенства для $t \in [t', t''] \subset [0, t_k]$

$$\sigma_i(t) = 1, \dot{\sigma}_i(t) = 0, \dots, \sigma_i^{(n)}(t) = 0, i = \overline{1, m}.$$

Образует систему алгебраических уравнений

$$\begin{aligned} \dot{\sigma}_i(t) &= \mathbf{b}_i^T \dot{\psi} = -\mathbf{b}_i^T \mathbf{A}^T \psi(t) = 0; \\ \ddot{\sigma}_i(t) &= -\mathbf{b}_i^T \mathbf{A}^T \dot{\psi}(t) = \mathbf{b}_i^T \mathbf{A}^{T^2} \psi(t) = 0; \\ &\dots\dots\dots; \\ \sigma_i^{(n)}(t) &= (-1)^n \mathbf{b}_i^T \mathbf{A}^{T^n} \psi(t) = 0, \end{aligned} \quad (19.52)$$

где вектор $\psi(t)$ — решение сопряженной системы дифференциальных уравнений

$$\dot{\psi} = -\partial \bar{H} / \partial x = -\mathbf{A}^T \psi.$$

Обозначив матрицу управляемости (п. 3.13.2) через G_i систему (19.52) запишем в виде

$$G_i^T \mathbf{A}^T \psi(t) = 0, i = \overline{1, m}. \quad (19.53)$$

Нетривиальное решение $\psi(t) \neq 0$ линейной однородной системы алгебраических уравнений (19.53) существует только в том случае, если

$$\det G_i^T \mathbf{A}^T = 0, i = \overline{1, m},$$

откуда следуют необходимые условия существования вырожденности

$$\det G_i = 0, i = \overline{1, m} \text{ и (или) } \det A = 0.$$

Условие $\det A = 0$ означает, что матрица A имеет хотя бы одно собственное число, равное нулю, т.е. объект управления содержит одно или более интегрирующих звеньев.

Достаточные условия отсутствия вырожденности

$$\det G_i \neq 0, i = \overline{1, m} \text{ и } \det A \neq 0, \quad (19.54)$$

т.е. объект должен обладать свойством нормальности и матрица A — невырожденная.

Синтез системы второго порядка, оптимальной по критериям быстрой работы и расхода ресурсов. Объект управления в виде двух последовательно включенных интегрирующих звеньев задан уравнениями

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u, \quad |u| \leq 1 \quad (19.55)$$

с граничными условиями $x(0) = x_0, x(t_k) = 0$.

Требуется найти управление $u(t), u(x)$, доставляющее минимум критерию вида (19.44), (19.42)

$$J = \int_0^{t_k} (\lambda + |u|) dt.$$

Запишем гамильтониан для уравнений объекта и критерия

$$\tilde{H} = -\lambda - |u| + \psi_1 x_2 + \psi_2 u,$$

а также сопряженные уравнения и соответствующие им решения

$$\dot{\psi}_1 = 0, \quad \psi_1(t) = \psi_{10},$$

$$\dot{\psi}_2 = -\psi_1, \quad \psi_2(t) = -\psi_{10}t + \psi_{20}. \quad (19.55a)$$

Оптимальное управление имеет вид (рис. 19.9).

$$u(t) = \text{dez} \psi_2(t).$$

Запишем решения $\sigma(t)$ для $t = t_1$ и $t = t_2$ (см. рис. 19.9):

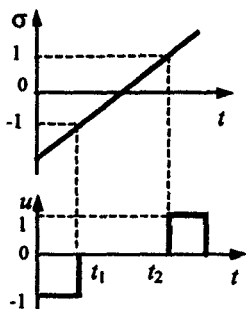


Рис. 19.9. График функций $\sigma(t)$ и $u(t)$

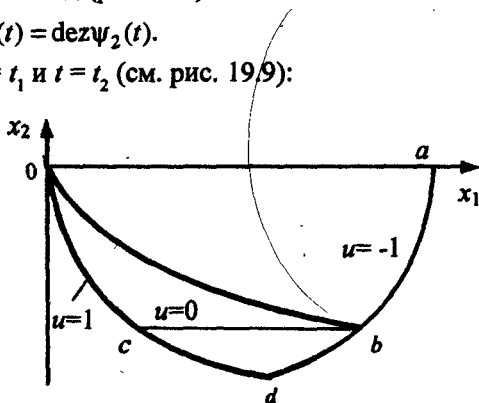


Рис. 19.10. Фазовые траектории системы, оптимальной по расходу топлива

$$\sigma(t_1) = -\psi_{10}t_1 + \psi_{20} = -1,$$

$$\sigma(t_2) = -\psi_{10}t_2 + \psi_{20} = 1,$$

откуда найдем

$$\psi_{10} = -2/\tau_1, \quad (19.56)$$

где $\tau_1 = t_2 - t_1$ — длительность интервала, на котором $u(t) = 0$.

В задаче со свободным конечным временем t_k функция $\tilde{H}(t) = 0$ для любого $t \in [0, t_k]$, поэтому из условия

$$\tilde{H}(t_1) = -\lambda - 1 + \psi_1 x_2(t_1) + 1 = 0$$

можно определить $\psi_{10} = \lambda/x_{21}$, где принято $x_{21} = x_2(t_1)$, и после сравнения с (19.56) получим

$$\tau_1 = -2x_{21}/\lambda, \quad (19.57)$$

Для решения задачи синтеза оптимального управления $u(x_1, x_2)$ применим метод фазовой плоскости. В отличие от задачи оптимального быстрогодействия при оптимизации расхода топлива необходимо найти две линии переключения (рис. 19.10). Линия переключения Od совпадает с линией переключения в задаче быстрогодействия, ее уравнение найдем тем же способом, что и § 19.4: разделим первое уравнение системы (19.55) на второе

$$dx_1/dx_2 = x_2/u,$$

и после интегрирования получим уравнение фазовой траектории Od при $u = 1$, проходящей через конечную точку $x(t_k) = 0$

$$x_1 = 0,5x_2^2. \quad (19.58)$$

Вторая линия переключения Ob есть геометрическое место точек переключения с $u(t) = -1$ на $u(t) = 0$. Для ее нахождения запишем решение системы (19.55) при $t \geq t_1$ и $u(t) = 0$

$$x_2(\tau) = x_{21}, \quad x_1(\tau) = x_{21}\tau + x_{11}, \quad (19.59)$$

где $\tau = t - t_1$, $x_{11} = x_1(t_1)$. Этому решению соответствует фазовая траектория bc на рис. 19.10. При $\tau = \tau_1$ будет достигнута точка c , в которой происходит переключение с $u(t) = 0$ на $u(t) = 1$. Координаты точки c (x_{1c} , x_{2c}) находятся при подстановке в решение (19.59) τ_1 согласно (19.57)

$$x_{2c} = x_2(\tau_1) = x_{21}, \quad x_{1c} = x_1(\tau_1) = -2x_{21}^2/\lambda + x_{11}.$$

Так как точка c находится также на траектории Od , то, приравняв x_{1c} решению (19.58) для точки c , получим уравнение

$$-2x_{21}^2/\lambda + x_{11} = x_{21}^2/2,$$

откуда, убрав вторые индексы, получим уравнение линии переключения 0b

$$x_1 = q_f x_2^2, \quad (19.60)$$

где $q_f = (4 + \lambda)/2\lambda$.

Зная линии переключения 0b и 0d, на их основе образуем две функции переключения

$$\begin{aligned} \mu_1 &= -x_1 - q_f |x_2| x_2, \\ \mu_2 &= -x_1 - |x_2| x_2 / 2, \end{aligned} \quad (19.60a)$$

приведенные в более общем виде, справедливым для любых x_{10}, x_{20} . Оптимальное управление с учетом расхода топлива формируется в соответствии с алгоритмом

$$u = \begin{cases} -1, & \text{при } \mu_1 < 0, \mu_2 < 0; \\ 0, & \text{при } \mu_1 \geq 0, \mu_2 < 0; \\ 1, & \text{при } \mu_1 > 0, \mu_2 > 0, \end{cases} \quad (19.60b)$$

который может быть реализован как аппаратным, так и программным способами.

19.5.2. Оптимизация по расходу энергии

Необходимо найти управление, доставляющее минимум критерию (19.41) с функцией $f_0(u)$ вида (19.43) или (19.43a) для линейного стационарного объекта (19.1).

Задача с фиксированным конечным временем без ограничения на управление. Для системы (19.1) и критерия (19.43) запишем функцию Гамильтона

$$\tilde{H} = -u^T R u + \psi^T A x + \psi^T B u.$$

В открытой области изменения $u(t)$ оптимальное управление находится из необходимого условия

$$\partial \tilde{H} / \partial u = -2R u + B^T \psi = 0,$$

откуда получим

$$u(t) = 0,5 R^{-1} B^T \psi(t). \quad (19.61)$$

Вычисление матрицы вторых производных $\partial^2 \tilde{H} / \partial u^2 = -2R < 0$ подтверждает, что при управлении (19.61) функция $\tilde{H}$ достигает максимума. Для нахождения $u(t)$ запишем сопряженную систему

$$\dot{\psi} = -\partial \tilde{H} / \partial x = -A^T \psi$$

и ее решение

$$\psi(t) = e^{-A^T t} \psi_0. \quad (19.62)$$

Подставив уравнение (19.61) в правую часть уравнения объекта (19.1) придем к системе

$$\dot{x} = Ax + 0,5BR^{-1}B^T \psi(t),$$

решение которой при $t = t_k$ с учетом (19.62) имеет вид

$$x(t_k) = e^{At_k} \left[x_0 + 0,5 \int_0^{t_k} e^{-A\tau} B R^{-1} B^T e^{-A^T \tau} d\tau \psi_0 \right]. \quad (19.63)$$

Так как время t_k задано, определенный интеграл может быть вычислен; обозначим результат интегрирования через матрицу V . Для случая фиксированных начального и конечного условий $x(0) = x_0$, $x(t_k) = x_k$ из решения (19.63) может быть вычислен вектор

$$\psi_0 = V^{-1}(e^{-At_k} x_k - x_0).$$

Отметим, что обратная матрица существует при условии полной управляемости объекта (19.1). Таким образом, для данного вида объекта и критерия с заданными конечным временем и граничными условиями удастся выразить начальное значение вектора ψ_0 через x_0 , x_k , t_k . Это позволяет найти согласно (19.61), (19.62) оптимальное управление $u(t)$ и решив уравнение (19.1), получить оптимальные переходные процессы $x(t)$.

Оптимизация по расходу энергии при ограничении на управление. Рассмотрим описание объекта в виде (19.45) и критерий (19.43а) при условии, что управляющие координаты ограничены по величине (19.2). Запишем гамильтониан

$$\tilde{H} = -r_1 u_1^2 - \dots - r_m u_m^2 + \psi^T A x + \psi^T b_1 u_1 + \dots + \psi^T b_m u_m.$$

Выделим часть функции $\tilde{H}$, явно зависящую от управлений u_i , $i = \overline{1, m}$,

$$H_i = -(r_i u_i^2 - \psi^T b_i u_i) - \dots - (r_m u_m^2 - \psi^T b_m u_m).$$

Аналогично задаче оптимизации по расходу топлива для достижения максимума H_i необходимо, чтобы каждое i -е слагаемое в круглых скобках было максимальным. Введем после деления на r_i функции

$$\sigma_i(t) = b_i^T \psi / 2r_i, \quad i = \overline{1, m}$$

и определим максимум путем выделения полного квадрата двучлена, т.е.

$$-(u_i - \sigma_i)^2 + \sigma_i^2 \rightarrow \max_{|u_i| \leq u}, \quad i = \overline{1, m}.$$

Запишем условия оптимальности $u_i(t)$ в зависимости от значения функции $\sigma_i(t)$, $i = \overline{1, m}$:

$$\begin{aligned} u_i(t) &= \sigma_i(t) && \text{при } |\sigma_i(t)| < \bar{u}_i, \\ u_i(t) &= \bar{u}_i \operatorname{sign} \sigma_i(t) && \text{при } |\sigma_i(t)| \geq \bar{u}_i \end{aligned}$$

или в сокращенной форме

$$u_i(t) = \bar{u}_i \operatorname{sat} \sigma_i(t), \quad i = \overline{1, m}. \quad (19.64)$$

Зависимость оптимального управления $u_i(t)$ от переменной $\sigma_i(t)$, $i = \overline{1, m}$ имеет вид характеристики усилительного звена с насыщением.

Отметим, что в отличие от задач быстродействия при оптимизации по расходу энергии отсутствует вырожденность, что следует из анализа решения (19.64).

Для наглядности на рис. 19.11 приведены характеристики нелинейных элементов, с помощью которых формировались оптимальные алгоритмы по критериям быстродействия, расхода топлива и энергии.

Синтез оптимальной системы второго порядка по критериям времени и расхода энергии. Выполним расчет системы, оптимальной по критерию

$$J = \int_0^{t_k} (\lambda + u^2) dt, \quad (19.65)$$

представляющему собой линейную комбинацию критериев времени и расхода энергии для объекта, представленного уравнениями (19.55). Для определения оптимального управления введем функцию Гамильтона

$$\tilde{H} = -\lambda - u^2 + \psi_1 x_2 + \psi_2 u. \quad (19.66)$$

Оптимальный закон управления имеет вид (рис. 19.12)

$$u(t) = \operatorname{sat} \sigma(t),$$

где $\sigma(t) = \psi_2(t)/2$, причем сопряженные уравнения и их решения совпадают с (19.55а)

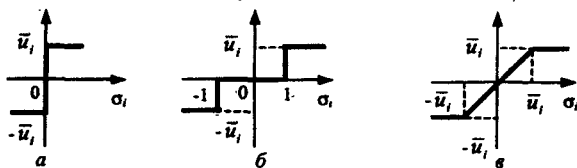


Рис. 19.11. Характеристики нелинейных элементов для формирования оптимального управления по критериям: а — времени, б — расхода топлива, в — расхода энергии

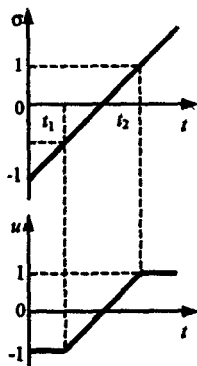


Рис. 19.12. График $\sigma(t)$ и $u(t)$

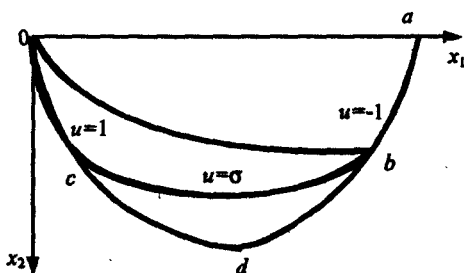


Рис. 19.13. Фазовые траектории системы, оптимальной по критериям времени и расхода энергии

Запишем условия, которым должна удовлетворять функция $\psi_2(t)$ при $t = t_1$ и $t = t_2$

$$\psi_2(t_1) = -\psi_{10}t_1 + \psi_{20} = -2,$$

$$\psi_2(t_2) = -\psi_{10}t_2 + \psi_{20} = 2,$$

откуда найдем

$$\psi_{10} = -4/\tau_1, \quad (19.67)$$

где $\tau_1 = t_2 - t_1$ — длительность интервала, на котором $u(t) = \sigma(t)$. В задаче со свободным конечным временем t_k функция $\tilde{H}(t) = 0$ для любого $t \in [0, t_k]$, поэтому из условия

$$\tilde{H}(t_1) = -\lambda - 1 + \psi_{10}x_2(t_1) + 2 = 0$$

можно определить $\psi_{10} = (\lambda - 1)/x_{21}$, где принято $x_{21} = x_2(t_1)$, и после сравнения с (19.67) получим при $\lambda > 1$

$$\tau_1 = -4x_{21}/(\lambda - 1). \quad (19.68)$$

Следуя способу решения, использованному при оптимизации по критериям быстродействия и расхода ресурсов для реализации алгоритма $u(t)$ (рис. 19.13) необходимо найти две линии переключения $0d$ и $0b$. Уравнение линии переключения $0d$ совпадает с (19.58), поэтому необходимо найти уравнение линии $0b$. Для этого запишем управление при $t \in [t_1, t_2]$

$$u(\tau) = d_1\tau + d_2, \quad (19.69)$$

где $\tau = t - t_1$, $d_1 = 2/\tau_1$, $d_2 = -1$. Выполнив интегрирование уравнений объекта (19.55), получим решение при $\tau = \tau_1$

$$x_2(\tau_1) = d_1 \tau_1^2 / 2 + d_2 \tau_1 + x_{21},$$

$$x_1(\tau_1) = d_1 \tau_1^3 / 6 + d_2 \tau_1^2 / 2 + x_{21} \tau_1 + x_{11}. \quad (19.70)$$

После подстановки в (19.70) выражения для τ_1 (19.68), а также d_1 , d_2 (19.69), и преобразования придем к соотношению

$$x_{1c} = g x_{21}^2 + x_{11}, \quad (19.71)$$

где коэффициент $g = (-4/3)(3\lambda - 1)/(\lambda - 1)^2$.

Так как моменту времени $\tau = \tau_1$ на фазовой плоскости соответствует точка с, принадлежащая одновременно траектории $0d$ при $u = 1$, то, выполнив действия, аналогичные п. 19.5.1, получим уравнение линии переключения $0b$

$$x_1 = q_e x_2^2,$$

где коэффициент $q_e = 0,5 - g$. Линия переключения $0b$ есть геометрическое место точек на фазовой плоскости, рис. 19.13, соответствующих переходу с управления $u = -1$ на *линейное во времени управление* $u(t) = \sigma(t)$. Функции переключения идентичны функциям (19.60а) после замены в функции μ_1 коэффициента q_f на q_e . Алгоритм оптимального управления соответствует (19.60б), отличаясь тем, что при $\mu_1 \geq 0$, $\mu_2 < 0$ управление $u(t) = \sigma(t)$.

Заметим, что этот алгоритм справедлив только при $\lambda > 1$, при котором $\psi_2(0) < -2$ и управление $u(0) = -1$ (см. рис. 19.12). При $0 < \lambda < 1$ оптимальное управление $|u(t)| < 1$, и, таким образом, необходимо реализовать только линейную функцию времени $u(t) = \sigma(t)$, $t \in [0, t_k]$.

19.5.3. Численный алгоритм оптимизации по расходу топлива и энергии для линейного объекта управления

В связи с тем, что для линейных систем известно аналитическое решение (п.2.7.2), при решении вариационных задач часто применяется метод преобразования к задаче *математического программирования*. Причем во многих случаях удается перейти к задачам линейного или квадратичного программирования, для решения которых разработаны эффективные численные методы оптимизации.

Применение линейного программирования для оптимизации по расходу топлива. Задача линейного программирования формулируется следующим образом. Требуется найти вектор x , доставляющий минимум линейной целевой функции

$$J = c^T x = c_1 x_1 + \dots + c_n x_n$$

при наличии линейных ограничений в виде как равенств, так и неравенств

$$Ax \begin{cases} \leq b_1, \\ = b_2, \\ \geq b_3. \end{cases} \quad (19.72)$$

и условия неотрицательности вектора $x \geq 0$.

Для перехода к такой задаче преобразуем критерии (19.41) и (19.42), заменив интеграл суммой, рассматривая частный случай *скалярного управления* ($m = 1$)

$$J = h \sum_{i=0}^{N-1} |u(i)|, \quad (19.73)$$

где $h = t_k / N$ — период дискретизации по времени; N — число шагов решения. От дифференциального уравнения (19.1) перейдем к разностному уравнению

$$x(i+1) = Fx(i) + gu(i), \quad (19.74)$$

где $F = e^{Ah}$, $g = \int_0^h e^{A\tau} b d\tau$.

Для задачи с заданными граничными условиями $x(0) = x_0$, $x(t_k) = x_k$ запишем решение уравнения (19.74) для $t_k = Nh$, $i = N-1$

$$x(N) = F^N x_0 + \sum_{i=1}^N F^{i-1} gu(N-i). \quad (19.75)$$

Так как $x(N) = x_k$, соотношение (19.75) есть линейное ограничение типа равенства. Целевая функция (19.73) является нелинейной, что не позволяет непосредственно применить линейное программирование. Для представления критерия (19.73) в линейной форме вводится преобразование

$$u(i) = u'(i) - u''(i), \quad (19.76)$$

где $u'(i) \geq 0$, $u''(i) \geq 0$, $i = \overline{0, N-1}$.

Функция цели (19.73) при этом представляется линейной относительно новых переменных $u'(i)$, $u''(i)$ [71]

$$J = h \sum_{i=0}^{N-1} [u'(i) + u''(i)], \quad (19.77)$$

а ограничения-неравенства на управление (19.2) имеют вид

$$u'(i) + u''(i) \leq \bar{u}, \quad i = \overline{0, N-1} \quad (19.78)$$

Таким образом, пришли к следующей задаче линейного программирования при оптимизации расхода топлива:

найти $u'(i)$, $u''(i)$, $i = \overline{0, N-1}$ при ограничениях в виде равенства (19.75) и неравенств (19.78), доставляющих минимальное значение показателю качества (19.77). Зная $u'(i)$, $u''(i)$, $i = \overline{0, N-1}$, по условиям (19.76) находится оптимальное управление $u(i)$, $i = \overline{0, N-1}$.

Приведем результаты численного расчета для уравнений объекта

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u, \quad |u| \leq 1 \quad (19.79)$$

и критерия
$$J = \int_0^{t_k} |u| dt$$

с заданными граничными условиями $x_1(0) = 1$, $x_2(0) = 0$, $x_1(t_k) = x_2(t_k) = 0$ и конечным временем $t_k = 2,5$. Число шагов на интервале $[0, 2,5]$ принято равным $N = 50$, при этом величина шага $h = 0,05$. Для решения задачи линейного программирования применялась функция $LP(c, A, b)$ из пакета MATLAB фирмы MathWorks, Inc. Оптимальные по расходу топлива переходные процессы приведены на рис. 19.14.

Применение квадратичного программирования для оптимизации по расходу энергии. Рассмотрим объект (19.1) и критерии (19.41), (19.43а) для случая скалярного управления ($m = 1$). Поскольку подын-

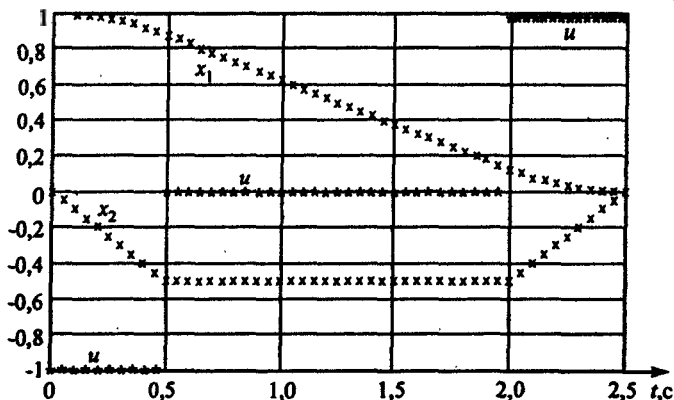


Рис. 19.14. Результаты численной оптимизации системы второго порядка по критерию расхода топлива

тегральное выражение в критерии есть квадратичная функция управления u , естественно свести исходную вариационную задачу оптимизации расхода энергии к квадратичному программированию.

Приведем формулировку задачи квадратичного программирования. Найти вектор x , доставляющий минимум функции цели

$$J = x^T H x + c^T x \quad (19.80)$$

при выполнении линейных ограничений (19.72) и условия $x \geq 0$.

От интегрального критерия перейдем к его дискретному аналогу

$$J = h \sum_{i=0}^{N-1} u^2(i), \quad (19.81)$$

где $h = t_k / N$. В отличие от задачи оптимизации расхода топлива показатель качества (19.81) соответствует стандартной функции цели (19.80). Таким образом, учитывая, что ограничения на управление $-u \leq u(i) \leq \bar{u}$, $i = \overline{0, N-1}$ являются линейными, исходная вариационная задача сведена к квадратичному программированию.

Пример решения задачи оптимизации по критерию расхода энергии выполнен для объекта (19.79) и критерия $J = \int_0^{t_k} u^2 dt$ с граничными условиями $x_1(0) = 1$, $x_2(0) = 0$, $x_1(t_k) = 0$, $x_2(t_k) = 0$ и конечным временем $t_k = 2,2$. Число шагов на интервале $[0, t_k]$ принято равным $N = 50$, при этом величина шага $h = 0,044$. Для решения задачи квадратичного программи-

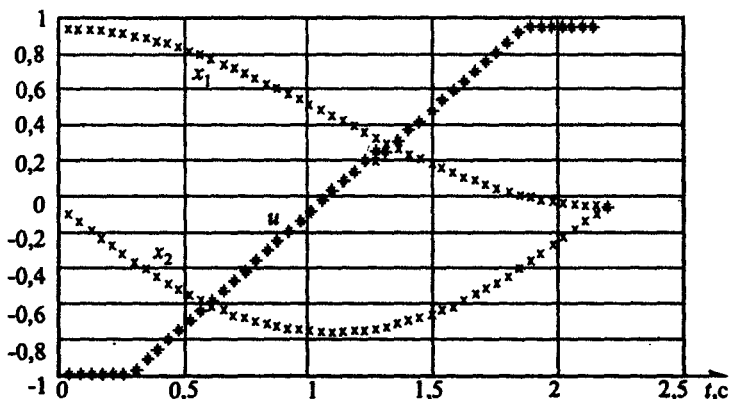


Рис. 19.15. Результаты численной оптимизации системы второго порядка по критерию расхода энергии

рования использовалась функция $QP(N, c, A, b)$ пакета MATLAB. Результаты численного расчета оптимальных процессов приведены на рис 19.15.

Численная оптимизация по комбинированному критерию. На базе программ расчета процессов, оптимальных по критериям расхода топлива или энергии при фиксированном конечном времени t_k , могут быть решены задачи оптимизации по комбинированным критериям

$$J' = \int_0^{t_k} (\lambda + |u|) dt \text{ или } J' = \int_0^{t_k} (\lambda + u^2) dt,$$

где $\lambda > 0$, время перехода t_k не задано.

Расчет выполняется следующим образом. Решается последовательность задач линейного или квадратичного программирования в соответствии с видом критерия с различными значениями t_k^i , $i = 1, 2, \dots$, задаваемыми с определенным шагом Δ . Для каждого t_k^i , $i = 1, 2, \dots$, в результате оптимизации при фиксированном конечном времени находится минимальное значение критерия расхода топлива или энергии $J(t_k^i)$, после чего вычисляется значение общего критерия $J' = \lambda t_k^i + J(t_k^i)$. Используя методы одномерного поиска (половинного деления и др.), находится значение t_k , при котором критерий J' достигает минимального значения. Заметим, что значения t_k^i , $i = 1, 2, \dots$ должны быть больше минимально возможного времени переходного процесса, соответствующего задаче оптимального быстрогодействия.

Глава 20

ОПТИМИЗАЦИЯ ДИНАМИЧЕСКИХ СИСТЕМ ПО КВАДРАТИЧНОМУ КРИТЕРИЮ

20.1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим линейный стационарный объект, описываемый уравнениями

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad (20.1)$$

где A, B, C — матрицы размером $n \times n$, $n \times m$, $p \times n$ соответственно; x — n -мерный вектор состояния; u — m -мерный вектор управления; y — p -мерный вектор выхода.

Качество работы системы управления задано квадратичным функционалом типа Больца

$$J = \int_0^{t_k} (x^T Q x + u^T R u) dt + x^T(t_k) F x(t_k), \quad (20.2)$$

Требуется найти управление $u(t, x)$, доставляющее минимум критерию (20.2) для объекта (20.1) при заданном начальном условии $x(0) = x_0$, фиксированном времени t_k и свободном $x(t_k)$.

Поясним назначение каждого из слагаемых в критерии (20.2). Первое слагаемое в подынтегральном выражении представляет собой квадратичную форму переменных $x_1, \dots, x_n$ и в случае диагональной матрицы Q с элементами $q_1, \dots, q_n$ имеет вид

$$x^T Q x = \sum_{i=1}^n q_i x_i^2.$$

Интеграл от этой функции имеет смысл интегральной квадратичной оценки, определяющей динамическую точность работы системы управления. При этом под x_i часто понимают отклонения переменных состояния от установившихся значений.

Второе слагаемое под интегралом также есть квадратичная функция управляющих координат $u_1, \dots, u_m$, приведем ее для частного случая диагональной матрицы R с элементами $r_1, \dots, r_m$

$$u^T R u = \sum_{i=1}^m r_i u_i^2.$$

Интеграл от этой функции характеризует расход энергии на управление за время t_k . Анализ показывает, что увеличение динамической точности системы происходит при одновременном увеличении интенсивности изменения управления и возрастании затрат энергии. Включение слагаемого $u^T R u$ в критерий приводит к косвенному ограничению $x(t)$, $y(t)$ и их производных. Рассмотрим, например, объект

$$\dot{x}_i = x_{i+1}, \quad \dot{x}_n = -a_1 x_1 - \dots - a_n x_n + u, \quad y = x_1, \quad i = \overline{1, n-1}$$

с переменными состояниями в виде $y, \dot{y}, \dots, y^{(n-1)}$ и критерий

$$J = \int_0^{\infty} (q_1 y^2 + u^2) dt.$$

После преобразования критерия с учетом уравнений объекта получим обобщенный квадратичный функционал

$$J = \int_0^{\infty} \left(\tau_0 y^2 + \tau_1 \dot{y}^2 + \dots + \tau_n y^{(n)^2} \right) dt,$$

используемый для параметрического синтеза систем управления (§ 4.9). Оптимальный процесс $y(t)$ при этом является решением линейного уравнения n -го порядка и, следовательно, регулятор может быть представлен в виде $u = -k_1 x_1 - \dots - k_n x_n$.

Таким образом, учет слагаемого $u^T R u$, введенного в работы Р.Е. Калмана, А.М. Летова, имеет принципиальное значение, поскольку для линейного объекта управления оказывается возможным получить закон управления как линейную функцию переменных состояния. Это в значительной мере определило большое внимание специалистов к данному классу оптимальных задач.

Рассмотрим последнее неинтегральное (терминальное) слагаемое в критерии (20.2), которое включается в функционал только в случае фиксированного времени t_k и свободного правого конца траектории $x(t_k)$. Введение терминальной составляющей дает возможность уменьшить значение $x(t)$ к моменту окончания управления t_k . Роль терминальной составляющей можно пояснить, преобразуя критерий (20.2) к виду

$$J = \int_0^{t_k} (x^T Q x + u^T R u + \delta(t - t_k) x^T(t) F x(t)) dt,$$

где $\delta(t - t_k)$ — дельта-функция — весовой коэффициент, принимающий бесконечно-большое значение в момент $t = t_k$.

В функционал могут быть включены выходные координаты

$$J = \int_0^{t_k} (y^T Q_1 y + u^T R u) dt + y^T(t_k) F_1 y(t_k). \quad (20.3)$$

После подстановки $y = Cx$ приходим к исходному критерию вида (20.2), где $Q = C^T Q_1 C$, $F = C^T F_1 C$.

Важное значение имеет учет влияния внешних воздействий в системе управления. В системах программного регулирования требуется, чтобы выход объекта как можно точнее воспроизводил задающий сигнал $g^*(t) = [g_1(t), \dots, g_p(t)]$. После введения ошибки воспроизведения $e = g - y$ критерий качества может быть сформирован аналогично (20.2)

$$J = \int_0^{t_k} (e^T Q e + u^T R u) dt + e^T(t_k) F e(t_k). \quad (20.4)$$

Кроме того, все системы работают в условиях влияния различных возмущений $\zeta^*(t) = [\zeta_1(t), \dots, \zeta_s(t)]$. Если возмущение $\zeta(t)$ детерминировано, то применяют один из рассмотренных выше критериев (20.2), (20.3), (20.4), учитывая уравнение объекта

$$\dot{x} = Ax + Bu + G\zeta, \quad y = Cx.$$

Если $\zeta(t)$ — случайное воздействие, тогда вместо критерия (20.2) вводят функционал

$$J = M \left[\int_0^{t_k} (\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt + \mathbf{x}^T(t_k) \mathbf{F} \mathbf{x}(t_k) \right],$$

где M — символ математического ожидания.

Функционал вида (20.2), а также (20.3), (20.4) косвенным образом за счет выбора весовых матриц $\mathbf{Q}$, $\mathbf{R}$ позволяет уменьшить значения управляющих координат в переходном процессе, не ограничивая, однако, значения скорости $\dot{\mathbf{u}}$, ускорения $\ddot{\mathbf{u}}$ и т.д. Более плавного изменения $\mathbf{u}(t)$ можно достичь также косвенным путем, добавив в критерий (20.2) слагаемые с $\dot{\mathbf{u}}$, $\ddot{\mathbf{u}}$... Пусть задан функционал для $m = 1$

$$J = \int_0^{t_k} (\mathbf{x}^T \mathbf{Q} \mathbf{x} + r \dot{\mathbf{u}}^2 + \eta \ddot{\mathbf{u}}^2) dt,$$

который приведем к стандартному виду (20.2) за счет дополнительной переменной состояния $x_{n+1} = \dot{\mathbf{u}}$. В результате этого запишем расширенную систему уравнений объекта

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x} + \mathbf{b} x_{n+1}, \quad \dot{x}_{n+1} = \ddot{\mathbf{u}}$$

и критерий

$$J = \int_0^{t_k} (\mathbf{x}^T \mathbf{Q} \mathbf{x} + r x_{n+1}^2 + \eta \ddot{\mathbf{u}}^2) dt.$$

При решении данной задачи будет получен закон управления относительно новой управляющей переменной $\ddot{\mathbf{u}} = \dot{\mathbf{u}} = \ddot{\mathbf{u}}(x_1, \dots, x_n, x_{n+1})$, так что при $\mathbf{x}(0) \neq 0$ реальное управление $\mathbf{u}(t)$ на входе объекта будет непрерывной функцией времени, начиная с $t = 0$.

Применение квадратичных критериев позволяет с единых позиций рассматривать синтез оптимальных систем различного класса: линейных стационарных и нестационарных, одномерных и многомерных, непрерывных и дискретных, с учетом задающих и возмущающих воздействий как детерминированных, так и случайных. Одна из главных трудностей в применении данных критериев связана с заданием весовых коэффициентов (элементов весовых матриц в (20.2)). Неопределенность в задании весовых коэффициентов обычно восполняется тем, что при расчете производится их варьирование с одновременным контролем получаемых показателей качества оптимальной системы, т.е. необходимо, как правило, неоднократное решение задачи оптимизации.

Особое значение при этом имеют вопросы связи весовых коэффициентов с известными показателями качества системы управления, такими, как полоса пропускания, время переходного процесса, точность в установившемся режиме и др. Многократное решение задачи синтеза в свою очередь эффективно лишь при малых затратах времени при однократном решении и, таким образом, важное значение имеет вычислительная сторона решения задачи оптимизации.

20.2. ДЕТЕРМИНИРОВАННЫЙ ОПТИМАЛЬНЫЙ РЕГУЛЯТОР С ФИКСИРОВАННЫМ ВРЕМЕНЕМ ПЕРЕХОДА ПРИ ПОЛНОЙ ИНФОРМАЦИИ О СОСТОЯНИИ ОБЪЕКТА

Пусть объект управления задан уравнениями

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{G}\zeta, \quad \mathbf{x}(0) = \mathbf{x}_0. \quad (20.5)$$

Требуется найти управления $\mathbf{u}(t)$, $\mathbf{x}(t)$, доставляющие минимум функционалу вида (20.2)

$$J = \int_0^t (\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt + \mathbf{x}^T(t_k) \mathbf{F} \mathbf{x}(t_k), \quad (20.6)$$

где $\mathbf{Q}$, $\mathbf{F}$ — положительно-полуопределенные матрицы, $\mathbf{R}$ — положительно-определенная матрица, все матрицы считаем симметрическими. В реальных условиях управляющие координаты ограничены

$$|u_i| \leq \bar{u}_i, \quad i = \overline{1, m}. \quad (20.7)$$

Выполним оптимизацию, применив принцип максимума. Учитывая особенности решения задачи для функционала Больца, запишем гамильтониан (18.7) для задачи (20.5), (20.6)

$$\tilde{H} = -\mathbf{x}^T \mathbf{Q} \mathbf{x} - \mathbf{u}^T \mathbf{R} \mathbf{u} + \boldsymbol{\psi}^T (\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{G}\zeta), \quad (20.8)$$

который требуется максимизировать по $\mathbf{u}$ при фиксированных $\boldsymbol{\psi}$, $\mathbf{x}$. Оптимальное управление в открытой области изменения $\mathbf{u}(t)$ находится из необходимого условия экстремума

$$\partial \tilde{H} / \partial \mathbf{u} = -2\mathbf{R}\mathbf{u} + \mathbf{B}^T \boldsymbol{\psi} = 0,$$

откуда

$$\mathbf{u}(t) = \mathbf{R}^{-1} \mathbf{B}^T \boldsymbol{\psi}(t) / 2. \quad (20.9)$$

Матрица вторых производных

$$\partial^2 \tilde{H} / \partial \mathbf{u}^2 = -2\mathbf{R}$$

отрицательно-определенная, т.е. при управлении (20.9) функция $\tilde{H}$ достигает максимума.

При определении оптимального управления с учетом ограничения (20.7) образуем функцию $\tilde{H}_1$, представляющую собой часть функции $\tilde{H}$ (20.8), зависящую от u

$$\tilde{H}_1 = -u^T R u + \psi^T B u.$$

Функция $\tilde{H}_1$ достигает максимума одновременно с $\tilde{H}$, поэтому оптимальное управление можно найти в результате решения задачи максимизации $\tilde{H}_1$ по u при $|u_i| \leq \bar{u}_i, i = \overline{1, m}$ (задачи квадратичного программирования). В данном случае можно найти решение путем преобразования $\tilde{H}_1$. С этой целью введем вектор-функцию

$$\sigma(t) = 0,5 R^{-1} B^T \psi(t)$$

и преобразуем функцию $\tilde{H}_1$ к следующему эквивалентному выражению

$$\tilde{H}_1 = -(u - \sigma)^T R (u - \sigma) + \sigma^T R \sigma.$$

Учитывая, что первое слагаемое не положительно, а последнее неотрицательно, получим условия максимума $\tilde{H}_1$ по u при $|u_i| \leq \bar{u}_i, i = \overline{1, m}$ в зависимости от значения функции $\sigma(t)$, записанные в покомпонентной форме:

$$\begin{aligned} u_i(t) &= \sigma_i(t) && \text{при } |\sigma_i(t)| < \bar{u}_i, \\ u_i(t) &= \bar{u}_i \operatorname{sign} \sigma_i(t) && \text{при } |\sigma_i(t)| \geq \bar{u}_i \end{aligned}$$

или в сокращенной форме:

$$u_i(t) = \bar{u}_i \operatorname{sat} \sigma_i(t), \quad i = \overline{1, m}. \quad (20.10)$$

Зависимость оптимального управления $u_i(t)$ от переменной $\sigma_i(t), i = \overline{1, m}$ имеет вид характеристики усилительного звена с насыщением.

Дальнейший расчет с учетом ограничения на основе алгоритма (20.10) связан с серьезными вычислительными трудностями. В связи с этим наиболее часто функционал (20.6) используют при расчете оптимального управления в *открытой* области изменения $u(t)$, рассматривая тем самым такие начальные отклонения $x(0)$, которые не приводят к выходу $u_i(t)$ на границу допустимой области. Однако после нахождения оптимального управления необходимо выполнить анализ характеристик системы с учетом реальных сигналов и ограничений (20.7).

Продолжим решение задачи, ориентируясь далее на уравнение (20.9). Запишем уравнение для вспомогательного вектора $\psi(t)$

$$\dot{\psi} = -\partial \tilde{H} / \partial x = 2 Q x - A^T \psi, \quad (20.11)$$

где $\psi(t_k)$ должен удовлетворять условию трансверсальности (18.20)

$$\psi(t_k) = -\partial \mathbf{x}^T(t_k) \mathbf{F} \mathbf{x}(t_k) / \partial \mathbf{x} = -2\mathbf{F} \mathbf{x}(t_k). \quad (20.11a)$$

Составим общую систему, присоединив к (20.11) уравнение объекта (20.5) с управлением (20.9),

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A} \mathbf{x} + 0,5 \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \psi + \mathbf{G} \zeta \\ \dot{\psi} &= 2\mathbf{Q} \mathbf{x} - \mathbf{A}^T \psi. \end{aligned} \quad (20.12)$$

Приведем решение этой системы

$$\begin{bmatrix} \mathbf{x}(t) \\ \psi(t) \end{bmatrix} = e^{-\mathbf{A}_0 t} \begin{bmatrix} \mathbf{x}_0 \\ \psi_0 \end{bmatrix} + \int_0^t e^{-\mathbf{A}_0(t-\tau)} \mathbf{G}_0 \zeta(\tau) d\tau,$$

где

$$\mathbf{A}_0 = \begin{bmatrix} \mathbf{A} & 0,5 \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \\ 2\mathbf{Q} & -\mathbf{A}^T \end{bmatrix}, \quad \mathbf{G}_0' = \begin{bmatrix} \mathbf{G} \\ \mathbf{0} \end{bmatrix}.$$

Для конечного момента времени t_k получим

$$\begin{bmatrix} \mathbf{x}(t_k) \\ \psi(t_k) \end{bmatrix} = e^{-\mathbf{A}_0(t_k-t)} \begin{bmatrix} \mathbf{x}(t) \\ \psi(t) \end{bmatrix} + \int_t^{t_k} e^{-\mathbf{A}_0(t_k-\tau)} \mathbf{G}_0 \zeta(\tau) d\tau. \quad (20.13)$$

Введем обозначение матрицы перехода

$$\Phi(t_k, t) = e^{-\mathbf{A}_0(t_k-t)} = \begin{bmatrix} \Phi_{11}(t_k, t) & \Phi_{12}(t_k, t) \\ \Phi_{21}(t_k, t) & \Phi_{22}(t_k, t) \end{bmatrix},$$

где каждый блок $\Phi(\cdot)$ есть матрица $n \times n$. Обозначим результат интегрирования в решении (20.13) через $\mathbf{v}(t) = \text{col} \begin{bmatrix} \mathbf{v}_1(t) & \mathbf{v}_2(t) \end{bmatrix}$, где $\mathbf{v}_1(t)$ и $\mathbf{v}_2(t)$ — n -мерные вектор-функции.

После преобразования выражения (20.13), учитывая блочный вид матрицы $\Phi(t_k, t)$ и (20.11a), получим уравнения (опуская аргументы)

$$\mathbf{x}(t_k) = \Phi_{11} \mathbf{x} + \Phi_{12} \psi + \mathbf{v}_1,$$

$$\psi(t_k) = \Phi_{21} \mathbf{x} + \Phi_{22} \psi + \mathbf{v}_2 = -2\mathbf{F} \mathbf{x}(t_k),$$

где во втором уравнении учтено условие трансверсальности (20.11a). После умножения обеих частей первого уравнения на $2\mathbf{F}$ и последующего сложения уравнений выразим вектор $\psi(t)$ через $\mathbf{x}(t)$ и $\mathbf{v}_1(t)$, $\mathbf{v}_2(t)$

$$\psi(t) = -2\mathbf{P}(t) \mathbf{x}(t) - 2\mathbf{g}(t), \quad (20.14)$$

где $2\mathbf{P}(t_k, t) = (\Phi_{22} + 2\mathbf{F}\Phi_{12})^{-1}(2\mathbf{F}\Phi_{11} + \Phi_{21})$,

$$2\mathbf{g}(t_k, t) = (\Phi_{22} + 2\mathbf{F}\Phi_{12})^{-1}(2\mathbf{F}\mathbf{v}_1 + \mathbf{v}_2).$$

После подстановки соотношения (20.14) в условие (20.9) получим оптимальное управление как функцию вектора состояния $\mathbf{x}(t)$ и возмущения $\zeta(t)$ (косвенно через $\mathbf{g}(t)$):

$$\mathbf{u}(t) = -\mathbf{K}(t)\mathbf{x}(t) - \mathbf{L}_g\mathbf{g}(t), \quad (20.15)$$

где $\mathbf{K}(t) = \mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}(t)$, $\mathbf{L}_g = \mathbf{R}^{-1}\mathbf{B}^T$ — матрицы коэффициентов усиления оптимального регулятора по $\mathbf{x}(t)$ и $\mathbf{g}(t)$ соответственно. Отметим, что приведенный здесь метод вычисления $\mathbf{K}(t)$ и $\mathbf{g}(t)$ через матрицы перехода является достаточно трудоемким.

20.2.1. Уравнение Риккати

Рассмотрим другой способ определения матриц $\mathbf{P}(t)$ и $\mathbf{g}(t)$ путем решения уравнения Риккати. Найдем это уравнение, выполнив следующие действия. Продифференцируем обе части соотношения (20.14)

$$\dot{\psi} = -2\dot{\mathbf{P}}\mathbf{x} - 2\mathbf{P}\dot{\mathbf{x}} - 2\dot{\mathbf{g}}.$$

Вместо $\dot{\mathbf{x}}$ подставим правую часть уравнения объекта (20.5) с учетом управления (20.9)

$$\dot{\psi} = -2\dot{\mathbf{P}}\mathbf{x} - 2\mathbf{P}\mathbf{A}\mathbf{x} + 2\mathbf{PBR}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x} + 2\mathbf{PBR}^{-1}\mathbf{B}^T\mathbf{g} - 2\mathbf{P}\mathbf{G}\zeta - 2\dot{\mathbf{g}}. \quad (20.16)$$

Из уравнения (20.12) и равенства (20.14) получим

$$\dot{\psi} = 2\mathbf{Q}\mathbf{x} + 2\mathbf{A}^T\mathbf{P}\mathbf{x} + 2\mathbf{A}^T\mathbf{g}. \quad (20.17)$$

В результате приравнивания правых частей уравнений (20.16) и (20.17) придем к уравнениям

$$\dot{\mathbf{P}} = -\mathbf{P}\mathbf{A} - \mathbf{A}^T\mathbf{P} + \mathbf{PBR}^{-1}\mathbf{B}^T\mathbf{P} - \mathbf{Q}, \quad (20.18)$$

$$\dot{\mathbf{g}} = \mathbf{PBR}^{-1}\mathbf{B}^T\mathbf{g} - \mathbf{A}^T\mathbf{g} - \mathbf{P}\mathbf{G}\zeta.$$

Последнее уравнение преобразуем, учитывая (20.15) и симметричность $\mathbf{P}$ (см. далее п. 20.2.2)

$$\dot{\mathbf{g}} = -(\mathbf{A} - \mathbf{BK})^T\mathbf{g} - \mathbf{P}\mathbf{G}\zeta. \quad (20.19)$$

Найдем граничные условия, необходимые для решения уравнений (20.18), (20.19), используя соотношения (20.11a), (20.14)

$$\psi(t_k) = -2\mathbf{P}(t_k)\mathbf{x}(t_k) - 2\mathbf{g}(t_k), \quad \psi(t_k) = -2\mathbf{F}\mathbf{x}(t_k),$$

откуда

$$\mathbf{P}(t_k) = \mathbf{F}, \quad (20.20)$$

$$\mathbf{g}(t_k) = \mathbf{0}. \quad (20.21)$$

Уравнение (20.18) есть *матричное дифференциальное уравнение Риккати*, оно не зависит от возмущения $\zeta(t)$, следовательно, его решение $P(t)$ и матрица обратных связей $K(t)$ не зависят от $\zeta(t)$. Зная $P(t)$ и определив $K(t)$ согласно (20.15), можно найти решение $g(t)$ уравнения (20.19). Матричному уравнению Риккати соответствует система n^2 нелинейных дифференциальных уравнений. Уравнение (20.19) представляет собой систему n линейных дифференциальных уравнений; его решение выполняют только в случае $\zeta(t) \neq 0$.

Граничные условия (20.20), (20.21) заданы для конечного времени t_k . Поэтому уравнения (20.18), (20.19) решают в *обратном времени* $\tau = t_k - t$:

$$dP/d\tau = PA + A^T P - PBR^{-1}B^T P + Q, \quad (20.22)$$

$$dg/d\tau = (A - BK)^T + PG\zeta \quad (20.23)$$

с начальными условиями $P(0) = F$, $g(0) = 0$.

В уравнении оптимального регулятора (20.15) используются *обратные* связи по переменным состояниям $x_1, \dots, x_n$, а также *прямые* связи, зависящие от возмущающего воздействия $\zeta(t)$, т.е. оптимальная система по своему принципу работы является *комбинированной*. Заметим, что реализация связей по возмущению затруднена, так как требуется заранее знать точный вид $\zeta(t)$ и, кроме того, уравнение (20.19) решается не в прямом, а в обратном времени. В отдельных случаях для задачи с $t_k = \infty$ удастся найти вид связей $g(t)$, например, для ступенчатых возмущений. Применяется также подход, использующий волновую модель возмущения в виде дополнительной системы дифференциальных уравнений, решение которой соответствует реальному характеру изменения $\zeta(t)$. При таком подходе, однако, требуется применение устройства оценивания для восстановления дополнительных переменных состояния [84].

Рассмотрим подробнее особенности оптимального решения при *отсутствии возмущения* $\zeta(t) = 0$ и, следовательно, $g(t) = 0$. Запишем уравнение регулятора (20.15) в по координатной форме

$$u_1 = -k_{11}(t)x_1 - \dots - k_{1n}(t)x_n,$$

$$\dots\dots\dots$$

$$u_m = -k_{m1}(t)x_1 - \dots - k_{mn}(t)x_n,$$

где коэффициенты $k_{ij}(t)$ *нестационарны*, зависят от времени, хотя объект управления (20.5) является стационарным. Нестационарность $k_{ij}(t)$ обусловлена тем, что время переходного процесса задано конечным. Это усложняет задачу реализации регулятора, но при использовании ЭВМ во время решения уравнения Риккати (20.22) можно организовать за-

поминание матриц $K(t_i)$, $t_i = iT$, $i = N-1, \dots, 0$, где T — период дискретизации, $t_k = NT$. В процессе управления объектом в реальном времени осуществляется выборка матриц регулятора $K(t_i)$ в обратном порядке, т.е. $K(0), \dots, K(N-1)$.

Заметим, что при определении алгоритма управления не требуется полной управляемости объекта, так как управление происходит на конечном отрезке времени t_k при свободном $x(t_k)$.

20.2.2. Свойства решения уравнения Риккати

Анализ правой части уравнения (20.18) показывает, что его решение *существует* и является *единственным* при заданном граничном условии (20.20). Непосредственной проверкой можно установить *симметричность* матрицы $P(t)$, что позволяет уменьшить число уравнений до $n(n+1)/2$. Знание $P(t)$ дает возможность вычислить *минимальное* значение функционала (20.6) (опуская промежуточные преобразования) при $\zeta(t) = 0$

$$\min J = x_0^T P(t_k, 0) x_0. \quad (20.24)$$

Если взять нижний предел интегрирования (20.6) равным t , то получим минимальное значение критерия как функцию времени

$$\min J(t) = x^T(t) P(t_k, t) x(t), \quad t \in [0, t_k]. \quad (20.25)$$

Так как каждое из слагаемых в функционале (20.6) неотрицательно, то матрица $P(t)$ для любого $t \in [0, t_k]$ является *положительно-определенной*. Учитывая функцию (18.44) в методе динамического программирования, можно считать, что выражение (20.25) есть функция Беллмана для критерия (20.6).

20.3. ДЕТЕРМИНИРОВАННЫЙ РЕГУЛЯТОР С БЕСКОНЕЧНЫМ ВРЕМЕНЕМ ПЕРЕХОДА ПРИ ПОЛНОЙ ИНФОРМАЦИИ О СОСТОЯНИИ

Рассмотрим практически важный случай, когда верхний предел интегрирования в критерии (20.6) $t_k = \infty$ при $\zeta(t) = 0$

$$J = \int_0^{\infty} (x^T Q x + u^T R u) dt. \quad (20.26)$$

Введение $t_k = \infty$ дает основание утверждать, что оптимальное управление $u(x)$, если оно существует, должно обеспечить устойчивость оптимального решения, т.е. $x(t) \rightarrow 0$ при $t \rightarrow \infty$, $x(0) \neq 0$; в противном случае интеграл (20.6) будет расходящимся. Поэтому терминальное слагаемое, входящее в (20.6), в критерий (20.26) не включено.

При определении оптимального управления применяют два способа. Один способ использует свойство устойчивости оптимальной системы, другой основан на определении предела $P(t_k, t)$ при $t_k \rightarrow \infty$, т.е. на рассмотрении последовательности решений задачи с фиксированным временем t_k при увеличении t_k .

Рассмотрим кратко первый способ. Решение системы (20.12) при $\zeta(t) = 0$ может быть представлено в следующем виде:

$$\begin{aligned} x_j(t) &= \sum_{i=1}^n \Delta_j(\lambda_i) c_i e^{\lambda_i t} + \sum_{i=1}^n \Delta_j(\lambda_{n+i}) c_{n+i} e^{\lambda_{n+i} t}, \quad j = \overline{1, n}, \\ \psi_j(t) &= \sum_{i=1}^n \Delta_{n+j}(\lambda_i) c_i e^{\lambda_i t} + \sum_{i=1}^n \Delta_{n+j}(\lambda_{n+i}) c_{n+i} e^{\lambda_{n+i} t}, \quad j = \overline{1, n}, \end{aligned} \quad (20.27)$$

где $\lambda_1, \dots, \lambda_n, \lambda_{n+1}, \dots, \lambda_{2n}$ — собственные числа матрицы A_0 , т.е. корни характеристического многочлена системы (20.12)

$$N_0(s) = \det(sI - A_0) = \det \begin{bmatrix} sI - A & -BR^{-1}B^T/2 \\ -2Q & sI + A^T \end{bmatrix}. \quad (20.28)$$

Причем, n корней $\lambda_1, \dots, \lambda_n$ имеют $\operatorname{Re} \lambda_i < 0$, $i = \overline{1, n}$ а остальные корни $\lambda_{n+1}, \dots, \lambda_{2n}$ имеют $\operatorname{Re} \lambda_{n+i} > 0$, $i = \overline{1, n}$. В решении (20.27) $\Delta_j(\lambda_i)$, $\Delta_{n+j}(\lambda_i)$ — алгебраические дополнения, соответствующие элементам одной из строк матрицы $(sI - A_0)$. Так как оптимальная система должна быть устойчивой, то все слагаемые, содержащие корни $\lambda_{n+1}, \dots, \lambda_{2n}$, приравниваются нулю. После чего получим систему из $2n$ линейных алгебраических уравнений с n неизвестными $c_i e^{\lambda_i t}$, $i = \overline{1, n}$. Исключив $c_i e^{\lambda_i t}$, $i = \overline{1, n}$ из решений $x_i(t)$, $i = \overline{1, n}$ и подставив их в последние n уравнений для $\psi_i(t)$, $i = \overline{1, n}$, придем к зависимости

$$\psi(t) = -2P\mathbf{x}(t), \quad (20.29)$$

где в отличие от задачи (20.14) с фиксированным t_k P — постоянная матрица $n \times n$. Учитывая, что в этом случае $\dot{P}(t) = 0$, матрица P может быть определена из матричного алгебраического уравнения

$$PA + A^T P - PBR^{-1}B^T P + Q = 0, \quad (20.30)$$

которое по аналогии с (20.18) также называют уравнением Риккати. Минимальное значение критерия (20.26) равно $\mathbf{x}_0^T P \mathbf{x}_0$.

Такой же результат достигается во втором способе решения, а именно доказано, что при $t_k \rightarrow \infty$ последовательность $\{P(t_k, t)\}$ стремится к пределу в виде постоянной матрицы P , которая удовлетворяет уравнению (20.30).

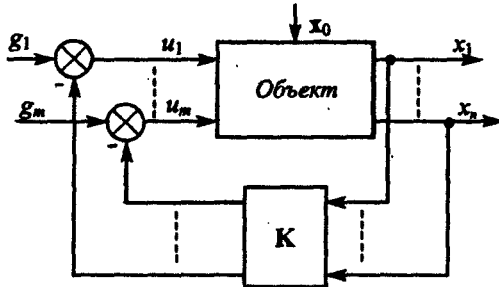


Рис. 20.1 Структурная схема системы, оптимальной по квадратичному критерию ($t_k = \infty$)

Из постоянства матрицы P следует, что матрица K , получающаяся после подстановки (20.29) в (20.9) тоже постоянна

$$u = -Kx, \quad K = R^{-1}B^T P. \quad (20.31)$$

Таким образом, при $t_k = \infty$ получен *стационарный* оптимальный регулятор (20.31) с *постоянной* матрицей коэффициентов K , что существенно упрощает реализацию такого регулятора (рис.20.1). Эта особенность решения обусловила большой интерес к таким системам и их достаточно широкое применение. Задающие воздействия $g_1(t), \dots, g_m(t)$ в уравнение регулятора (20.31) не входят; на рис. 20.1 они выполняют вспомогательную роль: с их помощью можно установить требуемые значения переменных состояния $x_1(t), \dots, x_n(t)$ объекта в установившемся режиме.

20.3.1. Устойчивость замкнутой системы с оптимальным регулятором

Запишем уравнение замкнутой оптимальной системы для объекта (20.1) и регулятора (20.31)

$$\dot{x} = (A - BK)x, \quad x(0) = x_0. \quad (20.32)$$

Оптимальный регулятор (20.31) гарантирует устойчивость замкнутой системы (20.32) при условии *полной управляемости* объекта и *положительной определенности* матрицы Q в критерии (20.26). Требование $Q > 0$ можно проверить, вычислив полную производную по времени от функции Ляпунова $S = x^T(t)Px(t)$ с учетом уравнений (20.32), (20.31), (20.30). Опуская преобразования, получим производную

$$dS/dt = -x^T Q x - u^T R u,$$

которая для устойчивости системы (20.32) должна быть отрицательной и, следовательно, Q — положительно-определенной матрицей.

Заметим, что реально условие $Q > 0$ при формировании функционала часто не выполняется. Можно показать, что в этом случае необходимо убедиться в полной наблюдаемости объекта (20.1) с фиктивным вектором выхода $z_f = Nx$, где $Q = N^T N$. Если объект не полностью управляем, то для достижения устойчивости замкнутой системы необходимо, чтобы неуправляемая подсистема объекта была устойчивой.

20.3.2. Учет постоянно действующих возмущений

Вернемся к расчету оптимального управления объектом (20.5) при $\zeta(t) \neq 0$. Так как $t_k = \infty$, необходимо учитывать так называемые *исчезающие* возмущения, для которых $\zeta(t) \rightarrow 0$ при $t_k \rightarrow \infty$, в противном случае значение интеграла (20.26) будет бесконечно большим. Например, при-

нимают $\zeta(t) = \sum_{i=1}^k c_i e^{\lambda_i t}$, где $\lambda_i < 0$. Причем, после определенной модификации такой класс возмущений позволяет найти управление и в случае типовых воздействий.

Приведем результаты расчета для практически важного случая *ступенчатого* возмущения $\zeta(t) = \zeta_0 1(t)$, базируясь на решении (20.19), полученном при фиксированном t_k . Ступенчатое воздействие $\zeta(t)$ представим в виде

$$\zeta(t) = \zeta_0 e^{\lambda t}, \quad \lambda < 0 \quad (20.33)$$

с любой степенью точности при выборе достаточно малого по величине значения λ .

Запишем уравнение (20.19)

$$\dot{g} = -V^T g - PG\zeta(t), \quad g(\infty) = 0, \quad (20.34)$$

где $V = A - BK$ — матрица замкнутой оптимальной системы с собственными числами в левой полуплоскости. Из решения уравнения (20.34)

$$g(t) = e^{-V^T t} g(0) - \int_0^t e^{-V^T (t-\tau)} PG\zeta(\tau) d\tau \quad (20.35)$$

определим $g(0)$ при котором $g(\infty) = 0$:

$$g(0) = \int_0^{\infty} e^{V^T \tau} PG\zeta(\tau) d\tau.$$

После подстановки $g(0)$ в решение (20.35) получим после преобразований

$$g(t) = \int_t^{\infty} e^{-V^T (t-\tau)} PG\zeta(\tau) d\tau. \quad (20.36)$$

Вычислив интеграл (20.36) для возмущения (20.33) при $\lambda \rightarrow 0$, найдем

$$g = V^{\tau^{-1}} P G \zeta_0$$

и уравнение оптимального регулятора (20.15)

$$u = -Kx + L\zeta_0,$$

где $L = R^{-1} B^{\tau} V^{\tau^{-1}} P G$ — постоянная матрица, что существенно упрощает задачу реализации регулятора (при условии, что $\zeta(t)$ можно измерить).

Рассмотрим еще один вид исчезающего возмущения $\zeta(t)$ — *дельта-функцию* $\delta(t)$. Из свойств линейной стационарной системы известно, что возмущение в виде $\delta(t)$ можно заменить эквивалентным вектором начальных условий $x_c(0)$. Это означает, что регулятор (20.31) обеспечивает оптимальность не только при свободном движении, вызванном $x(0) \neq 0$, но и при действии возмущения в виде $\delta(t)$. Очевидно, что независимость уравнения регулятора от возмущения $\zeta(t)$ сохранится и в случае, когда $\zeta(t)$ есть последовательность импульсов в виде $\delta(t - t_i)$, $0 \leq t_1 < t_2, \dots$

Поскольку сигнал в виде дельта-функции реально не существует, на практике приближенно можно считать, что приведенный здесь результат справедлив, если на объект действует последовательность импульсов малой длительности по сравнению с временем переходных процессов.

20.3.3. Вычисление матрицы стационарного оптимального регулятора

Для систем высокого порядка $n > 2$ аналитический расчет матрицы P в системе (20.30), а значит и K , практически невозможны. Поэтому широко применяются численные методы, ориентированные на ЭВМ.

Один из методов основан на свойстве предела матрицы $P(t_k, t)$, а именно $P(t_k, t) \rightarrow P$ при $t_k \rightarrow \infty$. Отсюда следует, что вместо алгебраического уравнения (20.30) решают дифференциальное уравнение Риккати (20.22) в обратном времени с $P(0) = 0$, останавливая расчет при достижении $P(\tau)$ установившегося значения с заданной точностью. Недостаток метода состоит в необходимости интегрирования системы нелинейных дифференциальных уравнений большой размерности, хотя в дальнейшем используется только одна точка решения.

Второй метод предназначен непосредственно для решения алгебраического уравнения (20.30), которое запишем в преобразованном виде, учитывая уравнение регулятора (20.31)

$$PA + A^{\tau} P - PBK + Q = 0.$$

Выбирают начальное приближение K_0 таким образом, чтобы замкнутая система с матрицей $A - BK_0$ была устойчивой (если объект управ-

ления устойчив при $K = 0$, то можно принять $K_0 = 0$). На каждой итерации решают матричное линейное алгебраическое уравнение относительно матрицы P_m , $m = 1, 2, \dots$

$$P_m A + A^T P_m - P_m B K_{m-1} + Q = 0,$$

$$K_m = R^{-1} B^T P_m,$$

Решение заканчивается при выполнении условия $\|P_m - P_{m-1}\| \leq \epsilon$, где ϵ — заданная точность решения.

Также известны методы, использующие представление (20.14) матрицы $P(t)$ через матрицу перехода системы (20.12) и преобразование матрицы A_0 этой системы к диагональной [39].

Пример. Заданы уравнение объекта управления

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u,$$

и критерий качества $J = \int_0^\infty (x_1^2 + u^2) dt$,

$$\text{т.е. } A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q = \begin{bmatrix} q_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad r = 1, \quad q_{11} = 1.$$

Запишем матричное уравнение Риккати

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} - \frac{1}{r} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} q_{11} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

После перемножения матриц получим

$$\begin{bmatrix} 0 & p_{11} \\ 0 & p_{12} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ p_{11} & p_{12} \end{bmatrix} - \frac{1}{r} \begin{bmatrix} p_{12}^2 & p_{12} p_{22} \\ p_{12} p_{22} & p_{22}^2 \end{bmatrix} + \begin{bmatrix} q_{11} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

В результате сложения матриц и приравнивания элементов матриц в обеих частях уравнения приходим к соотношениям

$$-p_{12}^2/r + q_{11} = 0,$$

$$p_{11} - p_{12} p_{22}/r = 0,$$

$$2p_{12} - p_{22}^2/r = 0.$$

Матрица P должна быть положительно-определенной, т.е. $p_{11} > 0$, $\det P > 0$, что позволяет получить решение в следующем виде:

$$p_{12} = \sqrt{r q_{11}} = 1, \quad p_{22} = \sqrt{2r} \sqrt{p_{12}} = \sqrt{2}, \quad p_{11} = p_{12} p_{22}/r = \sqrt{2}.$$

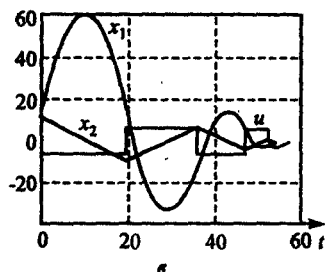
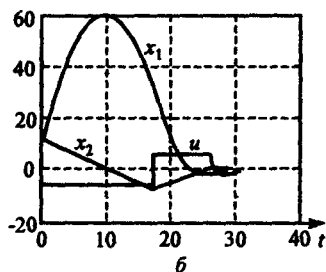
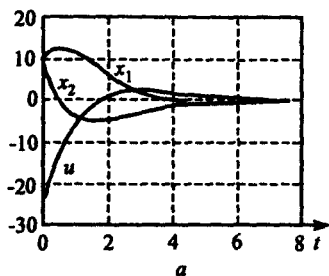


Рис. 20.2. Переходные процессы, оптимальные по квадратичному критерию: а — без ограничения $u(t)$; б — с ограничением $u(t)$; в — при $u(t) = -\text{sat } Kx$

В соответствии с (20.31) определим коэффициенты оптимального регулятора

$$K = [p_{12} \ p_{22}] / r = [1 \ \sqrt{2}],$$

т.е. $u = -x_1 - 1,414x_2$

Для анализа качества оптимальной системы запишем уравнение замкнутой системы

$$\dot{x} = (A - BK)x$$

и характеристический полином $s^2 + 1,414s + 1$; его корни равны $\lambda_{1,2} = -0,707 \pm j0,707$, коэффициент демпфирования $\xi = 0,707$. Переходные процессы в оптимальной системе при начальном условии $x(0) = [10 \ 10]$ приведены на рис. 20.2, а.

Рассмотрим влияние ограничения на управление $|u| \leq 1$. Аналитическое решение задачи оптимизации в этом случае невозможно, в связи с чем оптимальные процессы получены численным методом последовательных приближений (п. 18.2.4) и представлены на рис. 20.2, б. Ввиду трудности нахождения оптимальных процессов с учетом ограничения управления на практике часто применяют более простой в реализации линейный с насыщением регулятор

$$u = -\text{sat} Kx = \text{sat}(-x_1 - 1,414x_2).$$

Но, как показывает рис. 20.2, в, переходные процессы в рассматриваемой системе с таким регулятором имеют худшие показатели по сравнению с оптимальными (см. рис. 20.2, б). Отметим, что отличие в процессах зависит от начальных условий, возрастая с увеличением $x(0)$ (степени насыщения).

Для объекта со скалярным управлением ($m = 1$)

$$\dot{x} = Ax + bu \quad (20.37)$$

и критерия

$$J = \int_0^{\infty} (\mathbf{x}^T \mathbf{Q} \mathbf{x} + r u^2) dt \quad (20.38)$$

могут быть применены более простые способы вычисления коэффициентов оптимального регулятора $k_1, \dots, k_n$. Приведем один из них, состоящий из следующих действий.

Вначале вычисляют $2n$ корней характеристического многочлена $N_0(s)$ (20.28). Затем из этих корней отбирают n корней λ_i , имеющих $\operatorname{Re} \lambda_i < 0$, $i = \overline{1, n}$, которым соответствует характеристический полином устойчивой замкнутой оптимальной системы

$$N_1(s) = \prod_{i=1}^n (s - \lambda_i). \quad (20.39)$$

Для регулятора

$$u = -k_1 x_1 - \dots - k_n x_n = -\mathbf{k}^T \mathbf{x}$$

определяют характеристический многочлен замкнутой системы

$$N_2(s) = \det(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{k}^T).$$

Искомые коэффициенты $k_1, \dots, k_n$ находят путем сравнения многочленов $N_1(s)$ и $N_2(s)$, при этом потребуется решение линейной системы алгебраических уравнений относительно $k_1, \dots, k_n$. Таким образом, после вычисления корней λ_i , $i = \overline{1, n}$ определение $k_1, \dots, k_n$ свелось к решению известной задачи модального управления (п.4.3.2).

Если известна векторная передаточная функция объекта

$$\mathbf{W}(s) = \mathbf{x}(s)/u(s) = (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} = \mathbf{B}(s)/A(s), \quad (20.40)$$

где $A(s) = \det(s\mathbf{I} - \mathbf{A})$ — характеристический многочлен объекта, $\mathbf{B}(s) = [B_1(s) \dots B_n(s)]^T$ — вектор полиномов числителей $\mathbf{W}(s)$, то многочлены $N_0(s)$, $N_2(s)$ могут быть получены более простым путем согласно следующим выражениям:

$$N_0(s) = A(s)A(-s) + \mathbf{B}^T(-s)\mathbf{Q}\mathbf{B}(s)/r \quad (20.41)$$

или для случая диагональной матрицы $\mathbf{Q}$ с элементами q_i , $i = \overline{1, n}$

$$N_0(s) = A(s)A(-s) + \sum_{i=1}^n q_i B_i(-s)B_i(s)/r, \quad (20.41a)$$

$$N_2(s) = A(s) + \mathbf{k}^T \mathbf{B}(s) = A(s) + \sum_{i=1}^n k_i B_i(s). \quad (20.42)$$

20.3.4. Асимптотическое распределение корней

Так как при $t_k \rightarrow \infty$ оптимальный регулятор имеет постоянные коэффициенты, для выявления свойств оптимальной системы часто применяют известные методы анализа линейных систем управления.

Рассмотрим объект со скалярным управлением (20.37) и критерий (20.38). Корни полиномов (20.39) и (20.42) совпадают. Второе слагаемое в полиноме (20.41) преобразуем, выполнив его факторизацию

$$\mathbf{B}^T(-s)\mathbf{Q}\mathbf{B}(s)/r = k_e^2 B_e(-s)B_e(s), \quad k_e^2 = \alpha^2/r,$$

где $B_e(s) = \prod_{i=1}^v (s - \gamma_i)$, $\gamma_i < 0$, $i = \overline{1, v}$, α^2 — коэффициент, получающийся

в результате конкретных вычислений. Анализ корней полинома (20.41) при уменьшении $r \rightarrow 0$ на базе квадратичного корневого годографа дает следующее асимптотическое распределение устойчивых корней λ_i , $i = \overline{1, n}$ полинома $N_0(s)$, т.е. корней многочлена $N_2(s)$ (20.42).

При $r \rightarrow 0$ ($n - v$) корней многочлена $N_2(s)$ стремятся к корням стандартного полинома *Баттерворта* $N_B(s)$, а остальные v корней стремятся к корням γ_i полинома $B_e(s)$ [39]. Таким образом, при $r \rightarrow 0$ характеристический многочлен замкнутой системы примет вид $N_2(s) = N_B(s)B_e(s)$ причем все корни $N_B(s)$ располагаются на окружности радиуса $k_e^{1/(n-v)}$. Известно, что переходный процесс в замкнутой системе с передаточной функцией $1/N_B(s)$ имеет небольшое перерегулирование, что говорит о хороших динамических характеристиках оптимальной системы.

При исследовании многомерных систем ($m > 1$) необходимо анализировать характеристический многочлен

$$N_0(s) = A(s)A(-s)\det[\mathbf{I} + \mathbf{R}^{-1}\mathbf{W}^T(-s)\mathbf{Q}\mathbf{W}(s)/\rho]$$

при $\rho \rightarrow 0$, $\rho > 0$ [39].

20.3.5. Оптимальность в частотной области

Преобразуем многочлен (20.41) к виду

$$N_0(s)/A(s)A(-s) = 1 + \frac{1}{r}\mathbf{B}^T(-s)\mathbf{Q}\mathbf{B}(s)/A(s)A(-s). \quad (20.43)$$

Согласно (20.41) полином $N_0(s)$ содержит только четные степени s , поэтому он может быть представлен в виде произведения сопряженных

многочленов, т.е. $N_0(s) = N_2(s)N_2(-s)$. Следовательно, учитывая (20.40) и (20.42), получим соотношение

$$\frac{N_0(s)}{A(s)A(-s)} = \frac{N_2(s)N_2(-s)}{A(s)A(-s)} = (1 + \mathbf{k}^T \mathbf{W}(s))(1 + \mathbf{k}^T \mathbf{W}(-s)). \quad (20.44)$$

Из сравнения равенств (20.43) и (20.44) следует *условие оптимальности в частотной форме*

$$1 + \frac{1}{r} \mathbf{B}^T(-s) \mathbf{Q} \mathbf{B}(s) / A(-s) A(s) = (1 + \mathbf{k}^T \mathbf{W}(s))(1 + \mathbf{k}^T \mathbf{W}(-s)), \quad (20.45)$$

связывающее параметры регулятора с элементами функционала.

Если матрицу $\mathbf{Q}$ представить в виде $\mathbf{Q} = \mathbf{V}^T \mathbf{V}$, $\text{rang } \mathbf{V} = \text{rang } \mathbf{Q}$, то при условии полной наблюдаемости пары $(\mathbf{A}, \mathbf{V})$, соотношение (20.45) примет вид

$$1 + \frac{1}{r} \tilde{\mathbf{W}}^T(-s) \tilde{\mathbf{W}}(s) = (1 + \mathbf{k}^T \mathbf{W}(s))(1 + \mathbf{k}^T \mathbf{W}(-s)), \quad (20.45a)$$

где $\tilde{\mathbf{W}}^T(-s) = \mathbf{B}^T(-s) \mathbf{V}^T / A(-s)$, $\tilde{\mathbf{W}}(s) = \mathbf{V} \mathbf{B}(s) / A(s)$.

Заменив в (20.43) $N_0(s)$ на $N_2(s)N_2(-s)$, получим равенство при $s = j\omega$

$$\left| \frac{N_2(j\omega)}{A(j\omega)} \right|^2 = \left| 1 + \frac{\mathbf{B}^T(-j\omega) \mathbf{Q} \mathbf{B}(j\omega)}{r |A(j\omega)|^2} \right| = \left| 1 + \frac{\tilde{\mathbf{W}}^T(-j\omega) \tilde{\mathbf{W}}(j\omega)}{r |A(j\omega)|^2} \right|,$$

или

$$|N_2(j\omega) / A(j\omega)| \geq 1 \quad (20.46)$$

для $\omega = 0 + \infty$. Заметим, что $|N_2(s) / A(s)| \geq 1$ есть обратная передаточная функция ошибки замкнутой системы, считая выходом системы переменную $z = -\mathbf{k}^T \mathbf{x}$. Тогда условие (20.46) характеризует *чувствительность* оптимальной системы, а именно, такая система обладает свойством уменьшать влияние на систему возмущений и изменений параметров. Частотное условие оптимальности (20.45) дает возможность определить показатели качества оптимальной системы и разработать приближенные способы расчета коэффициентов регулятора. В [2] показано, что эквивалентная разомкнутая система с оптимальным регулятором имеет запас по фазе $\Delta\varphi_p \geq 60^\circ$ и по модулю $\Delta L \geq 2$; замкнутая система обладает показателем колебательности $M \leq 2$.

20.3.6. Критерий обобщенной работы

При решении задачи оптимизации по критериям (20.6) или (20.26) для линейного стационарного объекта требуется большой объем вычислений, связанный с решением нелинейного дифференциального

уравнения Риккати (20.22) или алгебраического (20.30). В [41] предложено видоизменить функционал (20.26), добавив в подынтегральное выражение слагаемое $u_{\alpha}^T R u_{\alpha}$, где $u_{\alpha} = -Kx$, $K = R^{-1}B^T P$ — сигнал на выходе оптимального регулятора, т.е. предполагается, что известен вид регулятора. Выражение $u_{\alpha}^T R u_{\alpha}$ иногда трактуется как обобщенная работа сигналов управления.

Запишем новый функционал

$$J = \int_0^{\infty} (x^T \tilde{Q} x + u^T R u) dt, \quad (20.47)$$

где $\tilde{Q} = Q + PBR^{-1}B^T P$.

Заменив в уравнении Риккати (20.30) матрицу Q на $\tilde{Q}$, придем к линейному уравнению типа Ляпунова

$$PA + A^T P + Q = 0, \quad (20.48)$$

решение которого проще, чем нелинейного уравнения (20.30): Если объект управления *устойчив*, Q — положительно-определенная матрица, то решение уравнения (20.48) есть положительно-определенная матрица P . Следовательно, замкнутая система с регулятором $u_{\alpha} = -R^{-1}B^T P x$ будет устойчивой.

Методика распространена на нейтральные и неустойчивые объекты, а также на другие виды линейных и нелинейных объектов управления.

20.4. СТОХАСТИЧЕСКИЙ НЕПРЕРЫВНЫЙ ОПТИМАЛЬНЫЙ РЕГУЛЯТОР ПРИ ПОЛНОЙ ИНФОРМАЦИИ О СОСТОЯНИИ

Пусть на входе объекта управления действует случайное возмущение в виде гауссовского белого шума $\zeta(t)$

$$\dot{x} = Ax + Bu + G\zeta, \quad x(0) = x(0) = x_0, \quad (20.49)$$

где $M[\zeta(t)] = 0$, $M[\zeta(t)\zeta^T(\tau)] = V\delta(t - \tau)$, V — матрица интенсивности белого шума, причем $\zeta(t)$ не коррелированно с x_0 .

Требуется найти управление $u(t, x)$, доставляющее минимум функционалу

$$J = M \left[\int_0^{t_k} (x^T Q x + u^T R u) dt + x^T(t_k) F x(t_k) \right] \quad (20.50)$$

для объекта (20.49) при фиксированном t_k .

Решение задачи оптимизации выполним методом стохастического динамического программирования. Запишем уравнение Беллмана (18.57) для объекта (20.49) и критерия (20.50)

$$\partial S / \partial t = - \min_{u(t)} \left\{ \mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u} + \left(\frac{\partial S}{\partial \mathbf{x}} \right)^T (\mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u}) + \frac{1}{2} \text{tr} \left[\mathbf{G} \mathbf{V} \mathbf{G}^T \frac{\partial^2 S}{\partial \mathbf{x}^2} \right] \right\}. \quad (20.51)$$

Найдем управление, обеспечивающее минимум выражения в фигурных скобках, из необходимого условия экстремума

$$\mathbf{u} = -0,5 \mathbf{R}^{-1} \mathbf{B}^T \partial S / \partial \mathbf{x}. \quad (20.52)$$

После подстановки (20.52) в уравнение (20.51) и преобразований получим уравнение в частных производных

$$\partial S / \partial t = - \mathbf{x}^T \mathbf{Q} \mathbf{x} + \frac{1}{4} \left(\frac{\partial S}{\partial \mathbf{x}} \right)^T \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \frac{\partial S}{\partial \mathbf{x}} - \left(\frac{\partial S}{\partial \mathbf{x}} \right)^T \mathbf{A} \mathbf{x} - \frac{1}{2} \text{tr} \left[\mathbf{G} \mathbf{V} \mathbf{G}^T \frac{\partial^2 S}{\partial \mathbf{x}^2} \right], \quad (20.53)$$

решение которого следует искать в виде

$$S(t, \mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x} + h(t), \quad (20.54)$$

где $\mathbf{P}(t)$ и $h(t)$ находятся в процессе следующих вычислений.

Найдем производные

$$\partial S / \partial \mathbf{x} = 2 \mathbf{P} \mathbf{x}, \quad \partial^2 S / \partial \mathbf{x}^2 = 2 \mathbf{P}, \quad \partial S / \partial t = \mathbf{x}^T \dot{\mathbf{P}} \mathbf{x} + \dot{h} \quad (20.55)$$

подставив которые в уравнение (20.53) после преобразований придем к уравнению Риккати (совпадающему с уравнением (20.18) детерминированной задачи)

$$\dot{\mathbf{P}} = -\mathbf{P} \mathbf{A} - \mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} - \mathbf{Q}, \quad \mathbf{P}(t_k) = \mathbf{F}$$

и уравнению

$$\dot{h} = -\text{tr}[\mathbf{G} \mathbf{V} \mathbf{G}^T \mathbf{P}(t)], \quad h(t_k) = 0. \quad (20.56)$$

Уравнение оптимального регулятора получим из условия (20.52) с учетом (20.55)

$$\mathbf{u} = -\mathbf{K}(t) \mathbf{x}, \quad \mathbf{K}(t) = \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P}(t). \quad (20.57)$$

Таким образом, получен важный результат, состоящий в том, что уравнения оптимальных регуляторов в стохастической и детерминированной задачах одинаковы и имеют одну и ту же матрицу $\mathbf{K}(t)$, не зависящую от возмущения $\zeta(t)$ (принцип *стохастической эквивалентности* в системах с полной информацией о векторе состояния).

Интуитивно этот результат можно объяснить, рассматривая белый шум как случайную последовательность некоррелированных импульсов в виде дельта-функции $\delta(t)$ со случайной «амплитудой» и бесконечно малым интервалом между импульсами. Тогда по аналогии с детерминированным возмущением типа $\delta(t)$ (п. 20.3.2) придем к полученному здесь результату.

Зная функцию Беллмана (20.54) определим

$$\min J = \mathbf{x}_0^T \mathbf{P}(0) \mathbf{x}_0 + h(0)$$

и, решив уравнение (20.56) в обратном времени, получим выражение для минимума критерия (20.50)

$$\min J = \mathbf{x}_0^T \mathbf{P}(0) \mathbf{x}_0 + \int_0^{t_k} \text{tr}[\mathbf{G} \mathbf{V} \mathbf{G}^T \mathbf{P}(t)] dt,$$

откуда следует, что возмущение $\zeta(t)$ увеличивает функционал на величину, соответствующую значению интеграла в последнем выражении.

Если система должна работать в течение длительного времени $t_k \rightarrow \infty$, то решают задачу оптимизации для объекта (20.49) и критерия

$$J = \lim_{t_k \rightarrow \infty} \frac{1}{t_k} M \left[\int_0^{t_k} (\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt \right]. \quad (20.58)$$

Также как в детерминированной задаче при $t_k \rightarrow \infty$, в результате решения стохастической задачи с критерием (20.58) получим постоянную матрицу $\mathbf{P}$ и оптимальный стационарный регулятор с постоянной матрицей регулятора $\mathbf{K}$. Минимальная величина критерия (20.58) равна

$$\min J = \mathbf{x}_0^T \mathbf{P} \mathbf{x}_0 + \text{tr}(\mathbf{G} \mathbf{V} \mathbf{G}^T \mathbf{P}).$$

20.5. ДИСКРЕТНЫЕ СИСТЕМЫ, ОПТИМАЛЬНЫЕ ПО КВАДРАТИЧНОМУ КРИТЕРИЮ

Рассмотрим вопрос построения дискретного регулятора, оптимального по квадратичному критерию качества. Потребность в этом возникает при использовании ЭВМ для управления различными объектами. Один из возможных путей состоит в применении непрерывного регулятора (20.15) с последующей его реализацией на ЭВМ при соответствующем выборе периода дискретизации по времени. Другой подход основан на том, что сначала производится дискретизация уравнения объекта, а затем рассчитывается оптимальный закон управления дискретным объектом. Именно этот случай рассмотрим ниже.

20.5.1. Дискретные детерминированные оптимальные системы

Пусть объект управления описывается разностным уравнением

$$\mathbf{x}_{n+1} = \Phi \mathbf{x}_n + \mathbf{D} \mathbf{u}_n, \quad \mathbf{x}(0) = \mathbf{x}_0. \quad (20.59)$$

Задан квадратичный критерий

$$J = \sum_{n=0}^{N-1} (\mathbf{x}_{n+1}^T \mathbf{Q} \mathbf{x}_{n+1} + \mathbf{u}_n^T \mathbf{R} \mathbf{u}_n) + \mathbf{x}_N^T \mathbf{F} \mathbf{x}_N. \quad (20.60)$$

Требуется определить такую последовательность управления $\mathbf{u}_0, \dots, \mathbf{u}_{N-1}$ объектом (20.59), которая доставляет минимум критерию (20.60) с фиксированным конечным временем $t_k = NT$ и свободным правым концом траектории $\mathbf{x}(t_k)$.

Для решения задачи воспользуемся методом динамического программирования. Запишем уравнение Беллмана (18.43) для критерия (20.60)

$$\begin{aligned} S_{N-j}(\mathbf{x}_{N-j}) = \min_{\mathbf{u}_{N-j}} [\mathbf{x}_{N-j+1}^T \mathbf{Q} \mathbf{x}_{N-j+1} + \mathbf{u}_{N-j}^T \mathbf{R} \mathbf{u}_{N-j} + \\ + \mathbf{x}_N^T \mathbf{F} \mathbf{x}_N + S_{N-j+1}(\mathbf{x}_{N-j+1})]. \end{aligned} \quad (20.61)$$

Необходимо учесть уравнение объекта (20.59) на j -м шаге

$$\mathbf{x}_{N-j+1} = \Phi \mathbf{x}_{N-j} + \mathbf{D} \mathbf{u}_{N-j}. \quad (20.62)$$

Учитывая линейность уравнения объекта и квадратичность функционала, зададим функцию Беллмана $S_{N-j+1}(\mathbf{x}_{N-j+1})$ в виде квадратичной функции

$$S_{N-j+1}(\mathbf{x}_{N-j+1}) = \mathbf{x}_{N-j+1}^T \mathbf{P}_{N-j+1} \mathbf{x}_{N-j+1}, \quad (20.63)$$

где $\mathbf{P}_{N-j+1}$ — симметрическая неотрицательно-определенная матрица, причем

$$\mathbf{P}_N = \mathbf{F}. \quad (20.64)$$

Подставим функцию (20.63) в уравнение (20.61) и положим

$$\mathbf{W}_{N-j+1} = \mathbf{P}_{N-j+1} + \mathbf{Q}, \quad (20.65)$$

в результате чего получим

$$S_{N-j}(\mathbf{x}_{N-j}) = \min_{\mathbf{u}_{N-j}} (\mathbf{x}_{N-j+1}^T \mathbf{W} \mathbf{x}_{N-j+1} + \mathbf{u}_{N-j}^T \mathbf{R} \mathbf{u}_{N-j}). \quad (20.66)$$

После подстановки решения (20.62) в (20.66) запишем уравнение, опуская временно аргументы,

$$S_{N-j}(\mathbf{x}_{N-j}) = \min_{\mathbf{u}} [(\Phi \mathbf{x} + \mathbf{D}\mathbf{u})^T \mathbf{W}(\Phi \mathbf{x} + \mathbf{D}\mathbf{u}) + \mathbf{u}^T \mathbf{R}\mathbf{u}] =$$

$$= \min_{\mathbf{u}} [\mathbf{x}^T \Phi^T \mathbf{W} \Phi \mathbf{x} + \mathbf{x}^T \Phi^T \mathbf{W} \mathbf{D} \mathbf{u} + \mathbf{u}^T \mathbf{D}^T \mathbf{W} \Phi \mathbf{x} + \mathbf{u}^T (\mathbf{D}^T \mathbf{W} \mathbf{D} + \mathbf{R}) \mathbf{u}].$$

Так как $\mathbf{W}$ симметрическая матрица, то второе слагаемое равно третьему, следовательно,

$$S_{N-j}(\mathbf{x}_{N-j}) = \min_{\mathbf{u}} (\mathbf{x}^T \Phi^T \mathbf{W} \Phi \mathbf{x} + 2\mathbf{x}^T \Phi^T \mathbf{W} \mathbf{D} \mathbf{u} + \mathbf{u}^T (\mathbf{D}^T \mathbf{W} \mathbf{D} + \mathbf{R}) \mathbf{u}). \quad (20.67)$$

Приравняем нулю градиент по $\mathbf{u}$ выражения под знаком минимума

$$\mathbf{D}^T \mathbf{W} \Phi \mathbf{x} + (\mathbf{D}^T \mathbf{W} \mathbf{D} + \mathbf{R}) \mathbf{u} = 0,$$

откуда находим оптимальное управление (восстанавливая индексы)

$$\mathbf{u}_{N-j} = -\mathbf{K}_{N-j} \mathbf{x}_{N-j}, \quad (20.68)$$

где

$$\mathbf{K}_{N-j} = (\mathbf{D}^T \mathbf{W}_{N-j+1} \mathbf{D} + \mathbf{R})^{-1} \mathbf{D}^T \mathbf{W}_{N-j+1} \Phi. \quad (20.69)$$

Зная $\mathbf{u}_{N-j}$, определим S_{N-j} в результате подстановки (20.68) с матрицей (20.69) в выражение (20.67) и преобразований

$$S_{N-j}(\mathbf{x}_{N-j}) = \mathbf{x}_{N-j}^T (\Phi^T \mathbf{W}_{N-j+1} \Phi - \Phi^T \mathbf{W}_{N-j+1} \mathbf{D} \mathbf{K}_{N-j}) \mathbf{x}_{N-j}.$$

После сравнения с функцией (20.63) придем к уравнению

$$\mathbf{P}_{N-j} = \Phi^T \mathbf{W}_{N-j+1} \Phi - \Phi^T \mathbf{W}_{N-j+1} \mathbf{D} \mathbf{K}_{N-j}. \quad (20.70)$$

Учитывая матрицы (20.69) и (20.65), от уравнения (20.70) перейдем к уравнению относительно $\mathbf{W}_{N-j}$

$$\mathbf{W}_{N-j} = \Phi^T \mathbf{W}_{N-j+1} [\mathbf{I} - \mathbf{D}(\mathbf{D}^T \mathbf{W}_{N-j+1} \mathbf{D} + \mathbf{R})^{-1} \mathbf{D}^T \mathbf{W}_{N-j+1}] \Phi + \mathbf{Q}. \quad (20.71)$$

Уравнение (20.71) называется матричным *разностным* уравнением Рикати, представляя собой дискретный аналог дифференциального уравнения для непрерывных систем. По условиям задачи $\mathbf{R}$ — симметрическая положительно-определенная матрица, следовательно, матрица $\mathbf{D}^T \mathbf{W} \mathbf{D} + \mathbf{R}$ также положительно-определенная и поэтому существует обратная матрица в уравнениях (20.69), (20.71). Заметим, что в отличие от непрерывной системы в дискретной возможно решение при $\mathbf{R} = \mathbf{0}$.

Выражения (20.65), (20.69), (20.70) представляют собой рекуррентные соотношения, позволяющие найти матрицу $\mathbf{K}_{N-j}$ регулятора и матрицу $\mathbf{P}_{N-j}$, $j = 1, \dots, N$, которые дают возможность определить значения критерия в соответствии с (20.63). Вычисления производят в *обратном времени*, начинаются они при $j = 1$ с заменой в выражении (20.69)

W_N на $F + Q$ согласно (20.64), (20.65) для получения K_{N-1} . Зная K_{N-1} и W_N , из (20.70) находим P_{N-1} . Затем идет второй шаг расчета: определяем W_{N-1} из (20.65), потом K_{N-2} , после чего P_{N-2} и т.д. до $j = N$, при котором найдем минимальное значение функционала $\min J = x_0^T P_0 x_0$.

Если не требуется знать на каждом шаге значение критерия, то выполняют решение уравнения Риккати (20.71) относительно матрицы W_{N-j} , $j = 1, \dots, N-1$ и вычисляют матрицу K_{N-j} (20.69). Матрицы K_i , $i = N-1, \dots, 0$ следует вычислить до начала работы системы управления и занести их в память ЭВМ. При включении системы в работу на каждом i -м шаге из памяти считывается соответствующая матрица в порядке, обратном расчетному, т.е. $K_0, \dots, K_{N-1}$.

Рассмотрим задачу управления объектом (20.59) при бесконечно большом времени перехода $t_k = NT = \infty$, считая, что в критерии (20.60) $F = 0$. При $N \rightarrow \infty$ решение разностного уравнения (20.71), как и в непрерывной системе, стремится к установившемуся, т.е. $W_{N-j} \rightarrow \bar{W}$ при $N \rightarrow \infty$, $j \rightarrow \infty$ и, следовательно, получим стационарный дискретный регулятор с постоянной матрицей K

$$u_i = -Kx_i. \quad (20.72)$$

Зависимости, необходимые для расчета матриц K , W примут вид

$$W = \Phi^T W [I - D(D^T W D + R)^{-1} D^T W] \Phi + Q, \quad (20.73)$$

$$K = (D^T W D + R)^{-1} D^T W \Phi. \quad (20.74)$$

Уравнение (20.73) есть матричное алгебраическое уравнение Риккати. После решения уравнения (20.73) вычисляют матрицу K . Минимальное значение критерия равно $\min J = x_0^T P x_0$, где $P = W - Q$. Для решения уравнения (20.73) разработаны различные численные методы [11, 84]. Отметим, что, также как в непрерывных системах, решение алгебраического уравнения (20.73) часто находят как установившееся решение в обратном времени разностного уравнения (20.71) с $W_N = Q$ при $N \rightarrow \infty$.

20.5.2. Оптимальные стохастические дискретные системы

Пусть задано уравнение дискретного объекта управления

$$x_{n+1} = \Phi x_n + D u_n + G \zeta_n, \quad x(0) = x_0, \quad (20.75)$$

где ζ_n — возмущение в виде дискретного белого шума, имеющее $M[\zeta_n] = 0$, $M[\zeta_n \zeta_j^T] = V \delta_{nj}$, где $\delta_{nj} = 1$ при $n = j$ и $\delta_{nj} = 0$ при $n \neq j$.

Требуется найти управляющую последовательность $u_0, \dots, u_{N-1}$, доставляющую минимум функционалу

$$J = M \left[\sum_{n=0}^{N-1} (x_{n+1}^T Q x_{n+1} + u_n^T R u_n) + x_N^T F x_N \right], \quad (20.76)$$

где предполагается фиксированным время перехода $t_k = NT$ и свободным конечное состояние x_N .

Решение задачи (20.75), (20.76) выполним методом стохастического динамического программирования. Оптимальное управление должно быть реализуемым $u_n = \eta(n, x_n)$, т.е. зависящим от текущего состояния объекта.

Рассмотрим последние j шагов управления, для которых запишем уравнение Беллмана

$$S_{N-j}(x_{N-j}) = \min_{u_{N-j}} M \left\{ \left[x_{N-j+1}^T Q x_{N-j+1} + u_{N-j}^T R u_{N-j} + S_{N-j+1}(x_{N-j+1}) \right] / x_{N-j} \right\}. \quad (20.77)$$

Допустим, что функция Беллмана имеет вид

$$S_{N-j+1}(x_{N-j+1}) = x_{N-j+1}^T P_{N-j+1} x_{N-j+1} + \alpha_{N-j+1}. \quad (20.78)$$

Подставив выражение (20.78) в уравнение (20.77), с учетом зависимости (20.65) получим

$$S_{N-j}(x_{N-j}) = \min_{u_{N-j}} M \left\{ \left[x_{N-j+1}^T W_{N-j+1} x_{N-j+1} + u_{N-j}^T R u_{N-j} + \alpha_{N-j+1} \right] / x_{N-j} \right\}. \quad (20.79)$$

Преобразуем первое слагаемое в (20.79), имея в виду уравнение объекта (20.75) и свойства условного математического ожидания (опуская индексы) [55, 73]

$$M \left\{ \left[x_{N-j+1}^T W_{N-j+1} x_{N-j+1} \right] / x_{N-j} \right\} = (\Phi x + Du)^T W (\Phi x + Du) + u^T D^T W D u + \text{tr}(G V G^T W).$$

После подстановки в (20.79) придем к уравнению

$$S_{N-j}(x_{N-j}) = \min_{u_{N-j}} \left[(\Phi x + Du)^T W (\Phi x + Du) + u^T (D^T W D + R) u + \text{tr}(G V G^T W) + \alpha \right]. \quad (20.80)$$

Сравнение уравнения (20.80) с уравнением детерминированной задачи (20.67) показывает совпадение слагаемых, содержащих управление u . Следовательно, оптимальное управление в данной стохастической задаче определяется выражениями (20.68), (20.69) и матрица регулятора K_{N-j} не зависит от возмущения ζ_n . От воздействия ζ_n зависят два последних слагаемых в уравнении (20.80) и значение функции (20.78), определяющей минимальное значение критерия.

Выполнив формальные преобразования при подстановке управления (20.68), (20.69) в уравнение (20.80), найдем дополнительно к (20.70), (20.71) уравнение для скалярной переменной α_n

$$\alpha_{N-j} = \text{tr}(\text{GVG}^T W_{N-j+1}) + \alpha_{N-j+1}, \quad \alpha_N = 0,$$

решение которого дает «добавку» к минимальному значению $x_0^T P_0 x_0$ в детерминированной системе, так что в стохастической задаче

$$\min J = x_0^T P_0 x_0 + \sum_{i=0}^{N-1} \text{tr}(\text{GVG}^T W_{i+1}). \quad (20.81)$$

Если время работы системы управления велико ($t_k = NT \rightarrow \infty$), то выполняют решение задачи оптимизации для критерия

$$J = \lim_{N \rightarrow \infty} \frac{1}{t_k} M \left[\sum_{n=0}^{N-1} (x_{n+1}^T Q x_{n+1} + u_n^T R u_n) \right]. \quad (20.82)$$

Также как в детерминированной задаче приходим к оптимальному регулятору с постоянной матрицей K (20.72) + (20.74). Минимальное значение критерия в этом случае равно $x_0^T P x_0 + \text{tr}(\text{GVG}^T W)$.

В настоящее время на практике широко применяется автоматизированный расчет параметров оптимальных регуляторов на базе пакета прикладных программ системы MATLAB [50].

Глава 21

ОПТИМАЛЬНЫЕ СИСТЕМЫ ПРИ НЕПОЛНОМ ИЗМЕРЕНИИ ВЕКТОРА СОСТОЯНИЯ

21.1. ОПТИМАЛЬНАЯ СИСТЕМА С ДЕТЕРМИНИРОВАННЫМ НАБЛЮДАТЕЛЕМ ПОЛНОГО ПОРЯДКА

Для реализации оптимального регулятора $u = -Kx$ требуется измерить все переменные состояния $x_1, \dots, x_n$. На практике оказывается, что часть координат вектора x нельзя использовать в уравнении регулятора

из-за отсутствия соответствующих датчиков, их низкой надежности или высокой стоимости. Иногда переменные состояния являются фиктивными, введенными в описание объекта для удобства вычислений. В то же время на объекте всегда имеются переменные, которые могут быть измерены, в частности выходные координаты, посредством каких-либо датчиков. Такие переменные образуют вектор измеряемых переменных

$$z = Hx.$$

Возникает задача восстановления (оценивания) координат вектора x по информации об уравнении объекта и вектора z . Выделим два способа, применяемых для восстановления x .

Первый из них состоит в том, чтобы выразить $x_1, \dots, x_n$ через $z_1, \dots, z_p$ и управляющие координаты $u_1, \dots, u_m$ (при условии, что измерения $z_1, \dots, z_p$ осуществляются точно, без помех). В этом случае в уравнение регулятора войдут производные от $z(t)$ и $u(t)$, что затрудняет применение данного способа.

Второй способ связан с построением специальных устройств (алгоритмов) оценивания (восстановления) x , называемых *наблюдающими устройствами* (НУ) или, кратко, *наблюдателями*. Эти устройства могут быть разработаны для детерминированных или стохастических задач в зависимости от условий работы объекта управления.

Пусть заданы уравнения объекта управления (ОУ) и p -мерного вектора измерения z

$$\dot{x} = Ax + Bu, \quad z = Hx, \quad p < n. \quad (21.1)$$

Необходимо по результатам измерения $z(t)$ и $u(t)$ найти оценку $\hat{x}$ вектора x так, чтобы $\hat{x}(t) \rightarrow x(t)$ при $t \rightarrow \infty$. Для построения наблюдателя используем уравнение

$$\dot{\hat{x}} = A\hat{x} + Bu + L(z - H\hat{x}). \quad (21.2)$$

Наблюдатель состоит из модели объекта (21.1), на вход модели вводится сигнал в виде ошибки (невязки) $z - H\hat{x}$ с матрицей усиления L , корректируя тем самым отклонение z от выхода модели $\hat{z} = H\hat{x}$.

Введем ошибку оценивания $e = x - \hat{x}$. Выполнив дифференцирование $\dot{e} = \dot{x} - \dot{\hat{x}}$, после подстановки правых частей уравнений (21.1), (21.2) и преобразований придем к уравнению наблюдателя для вектора ошибки

$$\dot{e} = (A - LH)e. \quad (21.3)$$

Матрица L должна выбираться так, чтобы система (21.3) была устойчивой, т.е. $e(t) \rightarrow 0$ при $t \rightarrow \infty$, $e(0) \neq 0$. Тогда $\hat{x}(t) \rightarrow x(t)$ при $t \rightarrow \infty$,

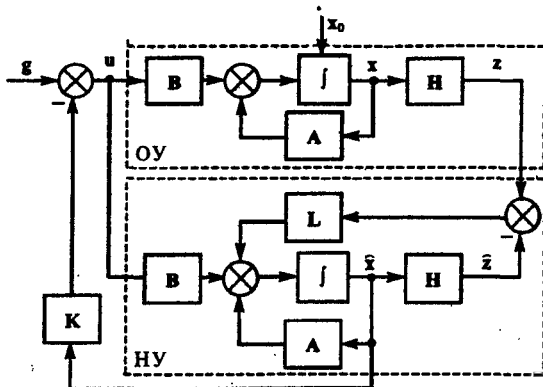


Рис. 21.1. Структурная схема оптимальной линейной системы с наблюдателем полного порядка

следовательно, наблюдатель (21.2) обеспечивает асимптотическое восстановление вектора состояния $x(t)$.

Если выполняется условие *наблюдаемости* системы (21.1), то выбором матрицы L можно достигнуть устойчивости наблюдателя и желаемого распределения корней его характеристического полинома $\det(sI - A + LH)$.

Оценка $\hat{x}(t)$ используется при формировании управления вместо истинного вектора $x(t)$

$$u = -K\hat{x}. \quad (21.4)$$

Так как наблюдатель (21.2) представляет собой динамическую систему, то он оказывает влияние на характеристики всей системы, рис. 21.1. Вектор задающих воздействий $g(t)$ не входит в уравнение регулятора (21.4), на рис. 21.1 он выполняет вспомогательную функцию, позволяя изменять режим работы ОУ.

Выясним это влияние путем анализа уравнения объекта (21.1) совместно с уравнениями наблюдателя (21.3) и регулятора (21.4)

$$\dot{x} = (A - BK)x + BKe,$$

$$\dot{e} = (A - LH)e$$

с характеристическим полиномом

$$N(s) = N_p(s)N_n(s),$$

где $N_p(s) = \det(sI - A + BK)$, $N_n(s) = \det(sI - A + LH)$.

Отсюда следует, что корни многочлена $N(s)$ включают в себя корни полинома $N_p(s)$ задачи оптимального регулирования и корни многочлена $N_n(s)$ наблюдателя. Это значительно упрощает расчет системы управления с наблюдателем, так как расчет матриц K и L выполняется *раздельно*, независимо друг от друга, за исключением этапа, связанного с назначением корней наблюдателя. Сначала вычисляется матрица K согласно (20.31), затем рассчитывается матрица L так, чтобы обеспечить заданные корни полинома $N_n(s)$. При этом корни $N_n(s)$, как отмечалось, должны быть расположены в левой полуплоскости корней и для увеличения скорости оценивания находиться левее корней $N_p(s)$.

Несмотря на наглядность расчета системы управления с наблюдателем, применение его вносит искажения в оптимальные процессы, зависящие от соотношения корней $N_p(s)$ и $N_n(s)$. В пределе влияние наблюдателя увеличением элементов матрицы L может быть сделано достаточно малым, но их увеличение реально допустимо до определенных значений, обусловленных действием шумов, помех в системе управления.

Аналогичным образом можно применить *дискретный детерминированный наблюдатель* для дискретного объекта (20.59) с вектором наблюдения $z_n = Hx_n$. Разностное уравнение наблюдателя в этом случае имеет вид

$$\hat{x}_{n+1} = \Phi \hat{x}_n + Du_n + L(z_n - H\hat{x}_n), \quad (21.5)$$

где матрица L выбирается так, чтобы обеспечить устойчивость наблюдателя и желаемое расположение собственных чисел матрицы $\Phi - LH$ внутри круга единичного радиуса.

21.2. ДЕТЕРМИНИРОВАННЫЙ НАБЛЮДАТЕЛЬ ПОНИЖЕННОГО ПОРЯДКА

Наблюдающие устройства полного порядка имеют определенную избыточность, так как независимо от размерности вектора измерения $z = Hx$ имеют максимальный порядок, равный n . Рассмотрим построение наблюдателя пониженного порядка $n-p$, где, как и ранее, p — размерность z .

Пусть имеем объект управления (21.1). Введем дополнительно к вектору z $(n-p)$ -мерный вектор $w^T = [w_1, \dots, w_{n-p}]$, связанный с вектором x соотношением

$$w = H_1 x,$$

где H_1 — произвольная матрица размером $(n-p) \times n$. Векторы z и w образуют n -мерный вектор

$$\tilde{z} = \begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} H \\ H_1 \end{bmatrix} x = D x, \quad (21.6)$$

Обычно H_1 выбирают простого вида, например, диагональной, но так, чтобы матрица D была неособенной, т.е. $\det D \neq 0$. Представим матрицу D^{-1} в блочном виде $D^{-1} = [L_1 \quad L_2]$, тогда из равенства (21.6) получим

$$x = D^{-1} \tilde{z} = L_1 z + L_2 w, \quad (21.7)$$

где матрицы L_1, L_2 размерностей $n \times p, n \times (n-p)$ соответственно.

В этом выражении вектор z известен, а вектор w должен быть восстановлен. Таким образом, задача построения наблюдателя сводится к нахождению оценки $\hat{w}$ вектора w и согласно (21.7) оценки

$$\hat{x} = L_1 z + L_2 \hat{w}. \quad (21.8)$$

С этой целью выполним дифференцирование $w(t)$ с учетом уравнения объекта (21.1)

$$\dot{w} = H_1 \dot{x} = H_1 A x + H_1 B u,$$

и после подстановки вместо x выражения (21.7) придем к уравнению

$$\dot{w} = H_1 A L_1 z + H_1 A L_2 w + H_1 B u. \quad (21.9)$$

Аналогичные действия выполним относительно вектора z :

$$\dot{z} = H A L_1 z + H A L_2 w + H B u, \quad (21.10)$$

откуда найдем

$$H A L_2 w = \dot{z} - H A L_1 z - H B u. \quad (21.11)$$

Из равенства (21.11) следует, что вектор $H A L_2 w$ может быть вычислен на основе измерений $z(t), u(t), \dot{z}(t)$. Следовательно, этот вектор совместно с уравнением (21.9) целесообразно использовать для нахождения оценки $\hat{w}$, применяя ту же схему, что и при формировании уравнения наблюдателя полного порядка (21.2).

Введем ошибку оценивания

$$e' = H A L_2 w - H A L_2 \hat{w} = \dot{z} - H A L_1 z - H B u - H A L_2 \hat{w}$$

и добавим ее с матрицей усиления M к правой части уравнения (21.9), которое используется как модель при формировании наблюдателя

$$\dot{\hat{w}} = H_1 A L_2 \hat{w} + H_1 A L_1 z + H_1 B u + M(\dot{z} - H A L_1 z - H B u - H A L_2 \hat{w}) \quad (21.12)$$

или

$$\begin{aligned} \dot{\hat{w}} = & (H_1 A L_2 - M H A L_2) \hat{w} + (H_1 A L_1 - M H A L_1) z + \\ & + (H_1 B - M H B) u + M \dot{z}. \end{aligned} \quad (21.12a)$$

Уравнение (21.12) или (21.12а) описывает наблюдатель *пониженного* порядка. Для того чтобы убедиться в этом, запишем уравнение для ошибки восстановления $e = w - \hat{w}$

$$\dot{e} = H_1 A L_2 w - H_1 A L_2 \hat{w} + M H A L_1 z + M H B u + M H A L_2 \hat{w} - M \dot{z}$$

и после подстановки вместо $\dot{z}$ правой части уравнения (21.10) получим уравнение ошибки

$$\dot{e} = (H_1 A L_2 - M H A L_2) e. \quad (21.13)$$

Если выбрать матрицу M так, чтобы система (21.13) была устойчивой, то $e(t) \rightarrow 0$ при $t \rightarrow \infty$, $e(0) \neq 0$ и, следовательно, $\hat{w}(t) \rightarrow w(t)$. Показано [39], что для полностью наблюдаемого объекта управления можно найти такую матрицу M , которая обеспечит любое желаемое расположение корней характеристического уравнения системы (21.13)

$$\det[sI - H_1 A L_2 + M H A L_2] = 0.$$

Реализация наблюдателя по уравнению (21.12а) требует дифференцирования переменных $z_1, \dots, z_p$, что часто выполнить затруднительно или можно осуществить только приближенно. Этот недостаток устраняется путем преобразования переменных

$$q = \hat{w} - Mz \text{ или } \hat{w} = q + Mz.$$

После дифференцирования $\dot{\hat{w}} = \dot{q} + M\dot{z}$ и подстановки в уравнение (21.12а) получим уравнение наблюдателя *пониженного* порядка, в котором исключена производная $\dot{z}$

$$\dot{q} = (H_1 A L_2 - M H A L_2) q + R z + T u, \quad (21.14)$$

где

$$R = H_1 A L_2 M + H_1 A L_1 - M H A L_1 - M H A L_2 M,$$

$$T = H_1 B - M H B.$$

Решение уравнения (21.14) позволяет определить векторы q , $\hat{w}$ и, следовательно, после подстановки в (21.8) найти оценку вектора x

$$\hat{x} = (L_1 + L_2 M) z + L_2 q.$$

Анализ замкнутой системы с регулятором $u = -K\hat{x}$ и наблюдателем (21.14) показывает, что характеристический многочлен такой системы имеет порядок, равный $2n - p$, причем, его корни есть объединение n корней оптимальной системы при условии точного измерения x и $(n - p)$ корней уравнения наблюдателя [39]. Таким образом, также как для наблюдателя полного порядка, справедливо свойство разделения.

21.3. ОПТИМАЛЬНЫЙ СТОХАСТИЧЕСКИЙ НЕПРЕРЫВНЫЙ НАБЛЮДАТЕЛЬ

Как отмечалось, при выборе матрицы L наблюдателя имеется определенный произвол, связанный с назначением корней полинома $N_{\lambda}(s)$ наблюдателя. Неопределенность в расчете матрицы L исчезает, если учесть более реальные условия работы объекта и канала измерения, а именно, влияние возмущающих воздействий.

Рассмотрим объект управления, на вход которого действует случайное возмущение $\zeta(t)$ (*шум состояния*)

$$\dot{x} = Ax + Bu + G\zeta. \quad (21.15)$$

В канале измерения также имеется возмущение $v(t)$ (*шум измерения*), т.е.

$$z = Hx + v(t). \quad (21.16)$$

Предположим, что $\zeta(t)$ и $v(t)$ представляют собой случайные процессы типа гауссовского белого шума, некоррелированные между собой, и имеют нулевые математические ожидания $M[\zeta(t)] = 0$, $M[v(t)] = 0$ и корреляционные матрицы

$$M[\zeta(t)\zeta^T(\tau)] = V_1\delta(t-\tau), \quad M[v(t)v^T(\tau)] = V_2\delta(t-\tau). \quad (21.17)$$

Симметрические матрицы интенсивностей V_1 и V_2 — неотрицательно-определенная и положительно-определенная соответственно. Элементы матрицы $V_2\delta(t-\tau)$, принадлежащие главной диагонали, представляют собой дисперсии измерительных шумов в соответствующих каналах измерения.

Начальное состояние объекта управления $x(t_0)$ есть случайная величина, некоррелированная с $\zeta(t)$ и $v(t)$, с известным средним значением $M[x(t_0)] = \bar{x}_0$, с заданной матрицей дисперсий

$$M[(x(t_0) - \bar{x}_0)(x(t_0) - \bar{x}_0)^T] = \Gamma_0. \quad (21.18)$$

Для оценивания $x(t)$ применим наблюдающее устройство той же структуры, что и в детерминированном случае (21.2)

$$\dot{\hat{x}} = A\hat{x} + Bu + L(z - H\hat{x}). \quad (21.19)$$

Введем ошибку восстановления $e = x - \hat{x}$. После дифференцирования $e(t)$ и преобразований с учетом уравнений (21.15), (21.16), (21.19) получим уравнение для ошибки

$$\dot{e} = (A - LH)e + G\zeta(t) + Lv(t). \quad (21.20)$$

Потребуем, чтобы наблюдатель давал *несмещенную* оценку вектора $x(t)$, т.е. должно выполняться условие $M[\hat{x}(t)] = M[x(t)]$, при этом $M[e(t)] =$

$= \bar{e}(t) = 0$. Условие $\bar{e}(t) = 0$ можно выполнить, полагая $\bar{e}(t_0) = 0$. Действительно, в уравнение (21.20) входят случайные возмущения $\zeta(t)$ и $v(t)$ с нулевыми средними значениями, поэтому $\bar{e}(t)$ вычисляется через матрицу перехода системы (21.20)

$$\bar{e}(t) = e^{(A-LH)t} \bar{e}(t_0).$$

Отсюда следует, что $\bar{e}(t) = 0$ при $\bar{e}(t_0) = 0$. Из равенства $\bar{e}(t) = 0$ находится начальное условие наблюдателя, при котором выполняется условие *несмещенности*,

$$\hat{x}(t_0) = \bar{x}_0.$$

Учитывая это равенство, преобразуем матрицу (21.18) к виду

$$M[(x(t_0) - \hat{x}_0)(x(t_0) - \hat{x}_0)^T] = M[e(t_0)e^T(t_0)] = \Gamma_0, \quad \hat{x}_0 = \bar{x}(t_0), \quad (21.21)$$

найдя таким образом матрицу дисперсий начального значения ошибки оценивания. Введем обозначение для матрицы дисперсий ошибки

$$M[e(t)e^T(t)] = \Gamma(t). \quad (21.22)$$

Зная уравнение (21.20), определим матрицу $\Gamma(t)$ как решение матричного уравнения

$$\dot{\Gamma} = (A - LH)\Gamma + \Gamma(A - LH)^T + GV_1G^T + LV_2L^T \quad (21.23)$$

с начальным условием (21.21), т.е. $\Gamma(t_0) = \Gamma_0$.

Запишем критерий оптимальности в виде математического ожидания квадратичной функции ошибки оценивания

$$J = M[e^T(t_k)Ne(t_k)] \rightarrow \min_L,$$

где N — симметрическая положительно-определенная матрица весовых коэффициентов, и преобразуем его, учитывая (21.22),

$$J = M[e^T(t_k)Ne(t_k)] = \text{tr}(\Gamma(t_k)N) \rightarrow \min_L. \quad (21.24)$$

Таким образом, необходимо найти такую матрицу $L(t)$, для которой решение $\Gamma(t)$ доставляет минимум критерию (21.24). Такой наблюдатель назовем *оптимальным*. Нахождение $L(t)$ есть вариационная задача оптимизации, которую решают различными способами, в том числе и с помощью принципа максимума. Действительно, матрицу $L(t)$ можно рассматривать как управления $l_{ij}(t)$, $i = \overline{1, n}$, $j = \overline{1, p}$, матрицу $\Gamma(t)$ как матрицу переменных состояния γ_{ij} , $i, j = \overline{1, n}$. В условиях данной оптимальной задачи время t_k фиксировано, а конечное состояние $\Gamma(t_k)$ свободно. Опуская преобразования, связанные с применением *матрично-*

го принципа максимума [66], приведем выражение для оптимальной матрицы коэффициентов усиления наблюдателя

$$L(t) = \Gamma(t)H^T V_2^{-1}, \quad (21.25)$$

где матрица $\Gamma(t)$ находится как решение матричного уравнения Риккати

$$\dot{\Gamma} = A\Gamma + \Gamma A^T - \Gamma H^T V_2^{-1} H \Gamma + G V_1 G^T, \quad \Gamma(t_0) = \Gamma_0, \quad (21.26)$$

получающегося после подстановки (21.25) в уравнение (21.23). Отметим, что матрица $L(t)$ — нестационарная, это вызвано тем, что оценивание осуществлялось на конечном отрезке времени $[t_0, t_k]$, где $t_0 = 0$.

Таким образом, определение матрицы $L(t)$ сводится к решению уравнения (21.26) в прямом времени (в отличие от задачи оптимального регулирования, где уравнение Риккати (20.18) решалось в обратном времени), что упрощает расчет и позволяет осуществить его в реальном времени. Из полученных результатов следует, что при заданных параметрах объекта, т.е. при известных матрицах A, B, H, G оптимальное решение $L(t)$ зависит только от характеристик шумов V_1, V_2 и начального значения матрицы дисперсий Γ_0 . Отметим, что получаемое решение не зависит от матрицы весовых коэффициентов N в критерии (21.24), что означает оптимальность оценки для каждой из координат $x_i(t)$, $i = \overline{1, n}$.

При оптимизации было принято, что уравнение наблюдателя (21.19) аналогично уравнению детерминированного наблюдателя (21.2). Р. Калман и Р. Бьюси получили именно такое уравнение (для нестационарного объекта) при решении задачи оптимальной фильтрации. Поэтому оптимальный наблюдатель состояния называют *фильтром* Калмана–Бьюси.

Приведем решение для оптимального фильтра в случае, когда шумы $\zeta(t)$ и $v(t)$ коррелированы, т.е.

$$M[\zeta(t)v^T(\tau)] = V_{12}\delta(t - \tau).$$

Введем обозначение $\tilde{V} = G V_{12}$. Матрица коэффициентов усиления наблюдателя определяется выражением

$$L(t) = (\Gamma(t)H^T + \tilde{V})V_2^{-1}, \quad (21.27)$$

дисперсионная матрица $\Gamma(t)$ находится путем решения уравнения Риккати [39]

$$\begin{aligned} \dot{\Gamma} = & (A - \tilde{V}V_2^{-1}H)\Gamma + \Gamma(A - \tilde{V}V_2^{-1}H)^T - \Gamma H^T V_2^{-1} H \Gamma + \\ & + G V_1 G^T - \tilde{V}V_2^{-1} \tilde{V}^T, \end{aligned} \quad (21.28)$$

где $\Gamma(t_0) = \Gamma_0$.

Также дополнительно рассмотрим задачу для измерений, включающих в себя *известные функции времени* $f(t)$

$$z = Hx + v(t) + f(t). \quad (21.29)$$

После введения нового вектора измерений

$$z_s = z - f(t)$$

уравнение наблюдателя примет вид

$$\dot{\hat{x}} = A\hat{x} + Bu + L(z_s - H\hat{x}), \quad (21.30)$$

где матрица L вычисляется таким же образом, что и при $f(t) = 0$.

Необходимость в решении последних задач возникает также при решении задачи оптимальной фильтрации в случае небелого (цветного) шума измерения или вырожденности матрицы V_2 .

21.3.1. Стационарный оптимальный наблюдатель

Ввиду сложности реализации матрицы L , зависящей от времени, важно найти условия, при которых эта матрица будет постоянной. Анализ показывает, что при увеличении интервала оценивания, т.е. $t_0 \rightarrow \infty$, матрица $L(t)$ стремится к установившемуся решению

$$L = \Gamma H^T V_2^{-1}, \quad (21.31)$$

где положительно определенная матрица Γ удовлетворяет матричному алгебраическому уравнению Риккати

$$A\Gamma + \Gamma A^T - \Gamma H^T V_2^{-1} H \Gamma + G V_1 G^T = 0. \quad (21.32)$$

Знание матрицы Γ дает возможность кроме вычисления матрицы коэффициентов усиления L определить согласно (21.22) *дисперсию ошибки оценивания*.

Поскольку время наблюдения равно бесконечности, наблюдатель должен быть асимптотически устойчивым. Матрица L (21.31) гарантирует устойчивость наблюдателя (собственные числа матрицы $A - LH$ располагаются в левой полуплоскости) при условии, что пара (A, H) полностью наблюдаема, а пара (A, G) — полностью управляема. В отличие от детерминированного наблюдателя (см. § 21.1), где корни характеристического уравнения могли быть назначены достаточно произвольно в левой полуплоскости корней, корни уравнения оптимального наблюдателя получаются вполне определенными, зависящими от интенсивностей шумов V_1, V_2 .

Для более наглядного выявления свойств оптимального наблюдателя получим *асимптотическое* распределение корней при изменении

V_1, V_2 . При этом будем опираться на свойства системы с оптимальным регулятором (см. п.п. 20.3.4, 20.3.5). С этой целью запишем следующую задачу оптимизации по квадратичному критерию:

уравнение фиктивного объекта

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}\tilde{u}, \quad (21.33)$$

где $\tilde{A} = A^T, \tilde{B} = H^T$,

критерий оптимальности

$$J = \int_0^{\infty} (\tilde{x}^T \tilde{Q} \tilde{x} + \tilde{u}^T \tilde{R} \tilde{u}) dt, \quad (21.34)$$

где $\tilde{Q} = G V_1 G^T, \tilde{R} = V_2$.

Для решения этой оптимизационной искусственно сконструированной задачи используем, как и в § 20.3, уравнение Риккати

$$P\tilde{A} + \tilde{A}^T P - P\tilde{B}\tilde{R}^{-1}\tilde{B}^T P + \tilde{Q} = 0,$$

которое после подстановки выражений для $\tilde{A}, \tilde{B}, \tilde{R}, \tilde{Q}$ совпадет с уравнением (21.32) оптимального стационарного наблюдателя (при этом $P = \Gamma$)

$$A\Gamma + \Gamma A^T - \Gamma H^T V_2^{-1} H \Gamma + G V_1 G^T = 0.$$

(Заметим, что аналогичным образом можно осуществить переход от задачи оптимального регулирования к задаче оптимального оценивания). Таким образом, задачи оптимального регулирования и оптимального наблюдения являются *дуальными* относительно друг друга. Это позволяет при соответствующей замене исходных данных переходить от решения одной задачи к другой.

После введения дуальной задачи оптимального регулирования можно использовать ранее полученные результаты в пп. 20.3.4, 20.3.5 при анализе асимптотических свойств оптимального наблюдателя. Рассмотрим случай *скалярного* измерения ($p = 1$) и скалярного шума состояния ($m = 1$), т.е. V_1, V_2 — заданные числа.

Запишем характеристический полином для задачи оптимального оценивания, аналогичный (20.41),

$$N(s) = A(s)A(-s) + B_n^T(-s)G V_1 G^T B_n(s)/V_2, \quad (21.35)$$

где $B_n(s)/A(s)$ — передаточная функция для системы (21.33):

$$\tilde{x}(s)/\tilde{u}(s) = B_n(s)/A(s) = (sI - A^T)^{-1} H^T$$

или

$$B_n^T(s)/A(s) = H(sI - A)^{-1}. \quad (21.36)$$

Преобразуем второе слагаемое в (21.35)

$$\mathbf{B}_n^T(-s)\mathbf{G}\mathbf{V}_1\mathbf{G}^T\mathbf{B}_n(s)/V_2 = \mathbf{V}_1\mathbf{G}^T\mathbf{B}_n(-s)\mathbf{B}_n^T(s)\mathbf{G}/V_2. \quad (21.37)$$

Запишем передаточные функции для системы (21.15), (21.16)

$$\begin{aligned} \mathbf{x}(s)/\zeta(s) &= (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{G} = \mathbf{B}_x(s)/A(s), \\ \mathbf{z}(s)/\zeta(s) &= \mathbf{H}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{G} = \mathbf{B}_z(s)/A(s). \end{aligned} \quad (21.38)$$

После сравнения (21.38) с (21.36) получим равенства

$$\mathbf{B}_n^T(s)\mathbf{G} = \mathbf{B}_z(s) \text{ и } \mathbf{G}^T\mathbf{B}_n(-s) = \mathbf{B}_z^T(-s),$$

в результате подстановки которых в (21.37), уравнение (21.35) примет вид

$$N(s) = A(s)A(-s) + \frac{V_1}{V_2}\mathbf{B}_z^T(-s)\mathbf{B}_z(s), \quad (21.39)$$

где $\mathbf{B}_z^T(-s)\mathbf{B}_z(s) = k_e^2 B_e(-s)B_e(s)$, $B_e(s) = \prod_{i=1}^{\beta} (s - \eta_i)$, $\eta_i < 0$, $i = \overline{1, \beta}$.

Как следует из выражения (21.39), распределение корней $N(s)$ зависит от отношения интенсивностей шумов V_1/V_2 . При $V_1/V_2 \rightarrow \infty$ (или $V_2 \rightarrow 0$ при постоянном V_1) асимптотическое распределение корней полинома (21.39) соответствует распределению корней полинома (20.41) (п.20.3.4). При значительной интенсивности шума измерения, т.е. $V_1/V_2 \rightarrow 0$, корни полинома наблюдателя (21.39) приближаются к корням полинома объекта $A(s)$. При этом коэффициент усиления наблюдателя $L \rightarrow 0$, а наблюдатель приближается к модели объекта, обеспечивая тем самым большую степень фильтрации. Если же измерительный шум мал, т.е. $V_1/V_2 \rightarrow \infty$, то $(n - \beta)$ корней полинома наблюдателя, стремясь к асимптотическому полиному Баттерворта, возрастают по величине (п.20.3.4) и $L \rightarrow \infty$. Это означает, что наблюдатель становится менее инерционным, с большей полосой пропускания.

21.3.2. Сингулярная (вырожденная) задача наблюдения

Необходимость в решении сингулярной задачи возникает в случае, когда $\det \mathbf{V}_2 = 0$, т.е. обратная матрица $\mathbf{V}_2^{-1}$ не существует и, следовательно, нельзя применить результаты, соответствующие (21.25), (21.26). Такая ситуация может возникнуть вследствие конкретных значений интенсивностей составляющих $v_i(t)$, $i = \overline{1, p}$ вектора шума измерения $\mathbf{v}(t)$, обусловленных условиями работы объекта и измерительных устройств. Из анализа (21.17) вырожденность всегда имеет место, если одна или

несколько переменных $z_i(t)$ не зависят от шума. В общем случае задачу с вырожденной матрицей V_2 можно преобразовать к виду [39], в котором выделены $\bar{p}$ переменных, свободных от шума,

$$z' = H'x \quad (21.40)$$

и $n - \bar{p}$ переменных, содержащих белый шум $\bar{v}$

$$\bar{z} = \bar{H}x + \bar{v}, \quad (21.41)$$

где $M[\bar{v}(t)\bar{v}'(\tau)] = \bar{V}_2\delta(t - \tau)$, $\det \bar{V}_2 \neq 0$. Шумы $\zeta_i(t)$ и $\bar{v}(t)$ не коррелированы, т.е. $M[\zeta_i(t)\bar{v}'(\tau)] = 0$.

При расчете наблюдателя выполним преобразования для вектора z' , аналогичные тем, которые были проведены в § 21.2 для детерминированного наблюдателя пониженного порядка. Дополним вектор z' до полной размерности, равной n , путем введения вектора $w' = H'_1x$ и выразим вектор x через z' и w' :

$$x = L'_1z' + L'_2w'. \quad (21.42)$$

После дифференцирования вектора w' аналогично (21.9) получим с учетом отличия дифференциального уравнения объекта (21.15) от (21.1)

$$\dot{w}' = H'_1AL'_1z' + H'_1AL'_2w' + H'_1Bu + H'_1G\zeta$$

$$\text{или} \quad \dot{w}' = A'w' + u' + G'\zeta, \quad (21.43)$$

где $A' = H'_1AL'_2$, $u' = H'_1AL'_1z' + H'_1Bu$, $G' = H'_1G$.

Так как вектор $z'(t)$ не зависит от белого шума, можно выполнить его дифференцирование

$$\dot{z}' = H'AL'_1z' + H'AL'_2w' + H'Bu + H'G\zeta \quad (21.44)$$

и тем самым получить дополнительную информацию, необходимую для построения наблюдателя.

Группируя известные функции $z'(t)$, $\dot{z}'(t)$, $u(t)$ в левой части (21.44) в соответствии с (21.29), (21.30), найдем соотношение для первой части нового вектора измерения

$$\bar{z}_1 = H'AL'_2w' + H'G\zeta, \quad (21.45)$$

где $\bar{z}_1 = \dot{z}' - H'AL'_1z' - H'Bu$.

Вторую часть вектора измерения получим, преобразуя выражение для вектора $\bar{z}$ (21.41)

$$\bar{z} = \bar{H}L'_1z' + \bar{H}L'_2w' + \bar{v}$$

$$\text{или} \quad \bar{z}_2 = \bar{H}L'_2w' + \bar{v}, \quad (21.46)$$

где $\bar{z}_2 = \bar{z} - \bar{H}L'_1z'$.

В результате объединения векторов $\bar{z}_1$ (21.45) и $\bar{z}_2$ (21.46) определим новый вектор измерения

$$\bar{z} = \begin{bmatrix} \bar{z}_1 \\ \bar{z}_2 \end{bmatrix} = \bar{H}w' + \bar{v}, \quad (21.47)$$

где $\bar{H} = \begin{bmatrix} H'AL'_2 \\ \bar{H}L'_2 \end{bmatrix}$, $\bar{v} = \begin{bmatrix} H'G\zeta \\ \bar{v} \end{bmatrix}$.

После выполненных преобразований получено дифференциальное уравнение (21.43) относительно переменной w' и уравнение измерения (21.47), в которое входит белый шум $\tilde{v}$. Причем шум состояния $\zeta(t)$ в уравнении (21.43) коррелирован с шумом измерения $\tilde{v}$ и $M[\zeta(t)\tilde{v}^T(\tau)] = V'_{12}\delta(t-\tau)$, где $V'_{12} = [V_1G'H^T0]$. Интенсивности шумов состояния $\zeta(t)$ и измерения $\tilde{v}$ равны соответственно

$$\bar{V}_1 = V_1, \quad \bar{V}_2 = M[\tilde{v}(t)\tilde{v}^T(t)] = \begin{bmatrix} H'GV_1G^TH^T & 0 \\ 0 & \bar{V}_2 \end{bmatrix}.$$

Согласно (21.27), (21.28) запишем уравнение оптимального наблюдателя $(n-p)$ -го порядка, считая матрицу $\bar{V}_2$ положительно определенной и вводя матрицу $\tilde{V} = G'V'_{12}$.

$$\dot{\hat{w}}' = A'\hat{w}' + u' + M'(\bar{z} - \bar{H}\hat{w}'), \quad \hat{w}'(0) \approx H'_1\bar{x}_0 \quad (21.48)$$

где $M' = (G'\bar{H}^T + \tilde{V})\bar{V}_2^{-1}$,

$$\dot{\Gamma}' = (A' - \tilde{V}\bar{V}_2^{-1}\bar{H})\Gamma' + \Gamma'(A' - \tilde{V}\bar{V}_2^{-1}\bar{H})^T -$$

$$-G'\bar{H}^B\bar{V}_2^{-1}\bar{H}\Gamma' + G'\tilde{V}_1G'^B - \tilde{V}\bar{V}_2^{-1}\tilde{V}^T, \quad \Gamma'(0) \approx H'_1\Gamma_0H_1^T. \quad (21.48a)$$

Приравнявая производную $\dot{\Gamma}'$ нулю, получим уравнения для стационарного оптимального наблюдателя.

Операцию дифференцирования $\dot{z}'(t)$ можно исключить, выполнив линейное преобразование переменных аналогично § 21.2. Для этого матрицу M' представим в блочном виде $M' = \begin{bmatrix} M'_1 & M'_2 \end{bmatrix}$ с учетом размерностей векторов $\bar{z}_1$, $\bar{z}_2$ и после подстановки в (21.48) придем к уравнению

$$\dot{\hat{w}}' = (A' - M'\bar{H})\hat{w}' + u' + M'_1\dot{z}' + M'_2\bar{z}_2 - M'_1(H'AL'_1\dot{z}' + H'Bu).$$

Выполним преобразование $q' = \hat{w}' - M'_1\dot{z}'$, после применения которого получим дифференциальное уравнение наблюдателя относительно вектора q' , не содержащее производной $\dot{z}'$.

Пример. Задан объект управления, описываемый уравнениями

$$\dot{x}_1 = -x_1 + x_2, \quad \dot{x}_2 = u + \zeta,$$

которым соответствуют матрицы

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Измеряемые переменные $z' = x_2$, $\bar{z} = x_1 + \bar{v}$, т. е. $H' = [01]$, $\bar{H} = [10]$. Шумы $\zeta(t)$ и $v(t)$ не коррелированы между собой и имеют интенсивности $V_1 = v_1$, $V_2 = v_2$; $\Gamma_0 = \text{diag}[\gamma_{01} \gamma_{02}]$.

В дополнение к точно измеряемой переменной z' введем переменную $w' = x_1 + x_2$, т. е. $H'_1 = [11]$, при этом матрица $D' = \begin{bmatrix} H'_1 \\ H' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ неособенная и, следовательно,

можно вычислить обратную матрицу $L' = D'^{-1} = [L'_1 \quad L'_2]$, где $L'_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $L'_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Запишем выражение (21.42) в по координатной форме $x_1 = -z' + w'$, $x_2 = z'$. После дифференцирования w' получим дифференциальное уравнение (21.43) $\dot{w}' = -w' + u' + \zeta$, где $u' = 2z' + u$. Дифференцирование z' и преобразование w' согласно (21.44), (21.45), (21.46) дают следующие соотношения для новых измеряемых переменных: $\bar{z}_1 = \zeta$,

$\bar{z}_2 = w' + \bar{v}$, где $\bar{z}_1 = z' - u$, $\bar{z}_2 = \bar{z} + z'$. Отсюда следует, что $\bar{H} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\bar{v} = \begin{bmatrix} \zeta \\ \bar{v} \end{bmatrix}$, что соответ-

ствует (21.47). Интенсивность шума $\bar{v}$ равна $\bar{V}_2 = \begin{bmatrix} v_1 & 0 \\ 0 & v_2 \end{bmatrix}$, причем $\det \bar{V}_2 \neq 0$. Шумы ζ

и $\bar{v}$ коррелированы, взаимная интенсивность $V'_{12} = [v_1 \quad 0]$, $\tilde{V} = [v_1 \quad 0]$. Запишем

уравнение наблюдателя $\dot{\hat{w}}' = -\hat{w}' + u' + M'(\bar{z} - H\hat{w}')$, где $M' = [m'_1 \quad m'_2]$, или

$\dot{\hat{w}}' = -(1 + m'_2)\hat{w}' + u' + m'_1\bar{z}_1 + m'_2\bar{z}_2$. С учетом соотношений для u' , $\bar{z}_1$, $\bar{z}_2$ получим

$$\dot{\hat{w}}' = -(1 + m'_2)\hat{w}' + (1 - m'_1)u + (2 + m'_2)z' + m'_1\dot{z}' + m'_2\bar{z}, \quad \hat{w}_0 \approx \bar{x}_{10} + \bar{x}_{20}.$$

Коэффициенты усиления наблюдателя определяются из выражения (21.48): $m'_1 = 1$, $m'_2 = \gamma/v_2$, где $\gamma'(t)$ находится путем решения уравнения Риккати (21.48а) $\dot{\gamma}' = -2\gamma' - \gamma'^2/v_2$ (при $\gamma' = \Gamma'$): $\gamma'(t) = -2v_2/(1 - d \cdot \exp(2t))$, где $d = (\gamma'_0 + 2v_2)/\gamma'_0$, $\gamma'_0 \approx \gamma_{01} + \gamma_{02}$. Зная γ' , $\hat{w}'$, определяются оценки $\hat{x}_1 = -z' + \hat{w}'$ и $\hat{x}_2 = z'$. Для исключения производной $\dot{z}'$ следует применить преобразование $q' = \hat{w}' - m'_1\dot{z}'$.

21.4. СТОХАСТИЧЕСКАЯ НЕПРЕРЫВНАЯ ОПТИМАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ПРИ НЕПОЛНОЙ ИНФОРМАЦИИ. ПРИНЦИП РАЗДЕЛЕНИЯ

Пусть объект управления и наблюдаемая переменная описываются уравнениями (21.15), (21.16) с теми же самыми характеристиками шумов состояния $\zeta(t)$ и измерения $v(t)$. Добавим к этим уравнениям соотношение для выходной (регулируемой) переменной

$$y = Cx. \quad (21.49)$$

Необходимо найти управление, использующее информацию о векторе наблюдения $z(t)$ (21.16), т.е. $u(t) = \varphi(t, z(t))$, при котором критерий

$$J = M \left[\int_0^{t_k} (y^T Q_1 y + u^T R u) dt + x^T(t_k) F x(t_k) \right] \quad (21.50)$$

достигает минимума. Учитывая зависимость (21.49), первое слагаемое в критерии (21.50) может быть записано в виде $x^T Q x$, где $Q = C^T Q_1 C$.

Решение задачи разбивается на два этапа, которые выполняются *раздельно*. На первом этапе осуществляют расчет оптимального стохастического наблюдателя (§ 21.3)

$$\dot{\hat{x}} = A\hat{x} + Bu + L(z - H\hat{x}), \quad (21.51)$$

$$L(t) = \Gamma(t)H^T V_2^{-1}, \quad (21.52)$$

$$\dot{\Gamma} = A\Gamma + \Gamma A^T - \Gamma H^T V_2^{-1} H \Gamma + G V_1 G^T, \quad \Gamma(0) = \Gamma_0. \quad (21.53)$$

Второй этап состоит в расчете матрицы $K(t)$ оптимального стохастического регулятора (20.57) при условии полного и точного измерения $x(t)$

$$K(t) = R^{-1} B^T P(t), \quad (21.54)$$

$$\dot{P} = -PA - A^T P + PBR^{-1}B^T P - Q, \quad P(t_k) = F. \quad (21.55)$$

Оптимальный регулятор в задаче (21.15), (21.16), (21.49), (21.50) описывается уравнениями, аналогичными (20.57) с заменой вектора $x(t)$ на его оценку $\hat{x}(t)$:

$$u = -K(t)\hat{x}. \quad (21.56)$$

Приведенные здесь условия оптимальности составляют *принцип разделения* (стохастической эквивалентности).

Рассмотрим кратко схему обоснования оптимального решения. Запишем ошибку оценивания $e(t) = x(t) - \hat{x}(t)$. Можно показать, что $e(t)$ и $\hat{x}(t)$ не коррелированы, т.е. $M[e(t)\hat{x}^T(t)] = 0$; кроме того, процесс $\tilde{\zeta} = z(t) - H\hat{x}(t)$ аналогично процессу $v(t) = z(t) - Hx(t)$ является белым шумом с интенсивностью V_2 [39]. Эти результаты дают возможность преобразовать критерий (21.50), выполнив следующие действия с учетом того, что $x(t) = e(t) + \hat{x}(t)$.

Определим

$$M[\mathbf{x}^T \mathbf{Q} \mathbf{x}] = M[(\mathbf{e} + \hat{\mathbf{x}})^T \mathbf{Q} (\mathbf{e} + \hat{\mathbf{x}})] = M[\mathbf{e}^T \mathbf{Q} \mathbf{e}] + M[\hat{\mathbf{x}}^T \mathbf{Q} \hat{\mathbf{x}}] + 2M[\mathbf{e}^T \mathbf{Q} \hat{\mathbf{x}}].$$

Так как $\mathbf{e}(t)$ и $\hat{\mathbf{x}}(t)$ не коррелированы и $M[\mathbf{e}^T(t) \mathbf{Q} \mathbf{e}(t)] = \text{tr}(\Gamma(t) \mathbf{Q})$,

$$M[\mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t)] = M[\hat{\mathbf{x}}^T(t) \mathbf{Q} \hat{\mathbf{x}}(t)] + \text{tr}(\Gamma(t) \mathbf{Q}).$$

Аналогичным образом преобразуем

$$M[\mathbf{x}^T(t_k) \mathbf{F} \mathbf{x}(t_k)] = M[\hat{\mathbf{x}}^T(t_k) \mathbf{F} \hat{\mathbf{x}}(t_k)] + \text{tr}(\Gamma(t_k) \mathbf{F}).$$

После подстановки в критерий (21.50) получим

$$J = M \left\{ \int_0^t (\hat{\mathbf{x}}^T \mathbf{Q} \hat{\mathbf{x}} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt + \hat{\mathbf{x}}^T(t_k) \mathbf{F} \hat{\mathbf{x}}(t_k) + \text{tr} \left[\int_0^t \Gamma(t) \mathbf{Q} dt + \Gamma(t_k) \mathbf{F} \right] \right\}, \quad (21.57)$$

где выражение в квадратных скобках не зависит от управления и его не надо учитывать при минимизации критерия. Таким образом, в результате преобразований пришли к критерию (21.57), содержащему вектор оценки $\hat{\mathbf{x}}$ в отличие от критерия (21.50). Если в качестве объекта взять наблюдатель (учитывая, что, как отмечалось, $\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t) = \tilde{\zeta}$ — белый шум с интенсивностью $\mathbf{V}_2$)

$$\dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}\tilde{\zeta}, \quad \hat{\mathbf{x}}(0) = \bar{\mathbf{x}}_0$$

и выполнить минимизацию (21.57), то придем к решению (21.56), (21.54), (21.55).

При увеличении времени регулирования $t_k \rightarrow \infty$ и времени оценивания $t_0 \rightarrow -\infty$ получим *стационарные* регулятор и наблюдатель с постоянными матрицами $\mathbf{K}$ и $\mathbf{L}$.

21.5. ОПТИМАЛЬНЫЙ ДИСКРЕТНЫЙ НАБЛЮДАТЕЛЬ

Рассмотрим задачу оптимального оценивания вектора состояния $\mathbf{x}_n$ для дискретного объекта, описываемого уравнением

$$\mathbf{x}_{n+1} = \Phi \mathbf{x}_n + \mathbf{D} \mathbf{u}_n + \mathbf{G} \zeta_n \quad (21.58)$$

и наблюдаемой переменной

$$\mathbf{z}_n = \mathbf{H} \mathbf{x}_n + \mathbf{v}_n, \quad (21.59)$$

Шум состояния ζ_n и шум измерения $\mathbf{v}_n$ — случайные гауссовские не коррелированные между собой последовательности с характеристиками $M[\zeta_n] = 0$, $M[\mathbf{v}_n] = 0$,

$$M[\zeta_n \zeta_n^T] = V_1 \delta_{nj}, \quad M[v_n v_j^T] = V_2 \delta_{nj}, \quad (21.60)$$

где $\delta_{nj} = 1$ при $n = j$ и $\delta_{nj} = 0$ при $n \neq j$.

Начальное состояние $x(0)$ есть случайная величина, имеющая $M[x(0)] = \bar{x}_0$ и

$$M[(x(0) - \bar{x}_0)(x(0) - \bar{x}_0)^T] = \Gamma_0,$$

причем, $x(0)$ не коррелировано с шумами ζ_n и v_n .

Допустим, что известна оценка $\hat{x}_n$ полученная по результатам измерения $z_1, \dots, z_n$. Значения оценки $\hat{x}_{n+1}^*$ найдем на основе прогноза на один шаг

$$\hat{x}_{n+1}^* = \Phi \hat{x}_n + D u_n. \quad (21.61)$$

Разность (невязка) между истинным z_n и предсказанным $z_n^* = H \hat{x}_n^*$ значениями добавляется при $t = (n+1)h$ к $\hat{x}_{n+1}^*$ для коррекции прогноза

$$\hat{x}_{n+1} = \hat{x}_{n+1}^* + L_{n+1}[z_{n+1} - H \hat{x}_{n+1}^*], \quad (21.62)$$

где L_{n+1} — матрица коэффициентов усиления наблюдателя.

Введем ошибку прогноза $\epsilon_n = x_n - \hat{x}_n^*$ и ошибку оценивания $e_n = x_n - \hat{x}_n$ и запишем разностные уравнения для ϵ_n и e_n , учитывая уравнения (21.58)–(21.62)

$$\epsilon_{n+1} = \Phi \epsilon_n + G \zeta_n, \quad (21.63)$$

$$e_{n+1} = (I - L_{n+1}H)(\Phi \epsilon_n + G \zeta_n) - L v_{n+1}. \quad (21.64)$$

Также как в непрерывном наблюдателе, для несмещенности $\hat{x}_n$ должно быть выполнено $M[e_n] = 0$ и $\hat{x}(0) = \bar{x}_0$.

Введем матрицу дисперсий

$$\Gamma_{n+1} = M[e_{n+1} e_{n+1}^T], \quad (21.65)$$

называемую апостериорной, так как она вычисляется после измерения z_{n+1} . Из уравнения (21.63) с учетом (21.60), (21.61) найдем априорную матрицу дисперсий

$$\tilde{\Gamma}_{n+1} = M[\epsilon_{n+1} \epsilon_{n+1}^T] = \Phi \Gamma_n \Phi^T + G V_1 G^T. \quad (21.66)$$

На основе уравнений (21.64), (21.66) получим для матрицы (21.65) (опуская аргументы) [39]

$$\begin{aligned} \Gamma_{n+1} &= \left\{ [M(I - LH)(\Phi e + G \zeta) - L v] [M(I - LH)(\Phi e + G \zeta) - L v]^T \right\} = \\ &= (I - L_{n+1}H) \tilde{\Gamma}_{n+1} (I - L_{n+1}H)^T + L_{n+1} V_2 L_{n+1}^T, \end{aligned} \quad (21.67)$$

так как $M[\zeta_n e_n^T] = 0$, $M[v_n e_n^T] = 0$, $M[\zeta_n v_n^T] = 0$.

Введем критерий оптимальности наблюдателя в виде суммы дисперсий ошибок оценивания или в векторной форме

$$J = M[e_n^T e_n] \rightarrow \min_L,$$

или, учитывая выражение (21.65),

$$J = \text{tr } \Gamma_n \rightarrow \min_L. \quad (21.68)$$

Таким образом, задача оценивания свелась к определению матрицы наблюдателя L_n в уравнении (21.67), доставляющей минимум критерию (21.68). Для решения данной задачи оптимизации проведем преобразование критерия (21.68), имея в виду уравнение (21.67)

$$J = \text{tr } \tilde{\Gamma}_{n+1} - 2\text{tr}(L_{n+1} \tilde{H} \tilde{\Gamma}_{n+1}) + \text{tr}[(L_{n+1}(\tilde{H} \Gamma_{n+1} \tilde{H}^T + V_2) L_{n+1}^T)]. \quad (21.69)$$

Применим правила дифференцирования следа матрицы к выражению (21.69) (опуская аргументы) [66]

$$\partial \text{tr}(L \tilde{H} \Gamma) / \partial L = \tilde{H} \Gamma^T,$$

$$\partial \text{tr}[L(\tilde{H} \Gamma \tilde{H}^T + V_2) L^T] / \partial L = 2L(\tilde{H} \Gamma \tilde{H}^T + V_2).$$

Применяя необходимое условие экстремума критерия (21.68)

$$\partial J / \partial L = -\tilde{H} \Gamma^T + L(\tilde{H} \Gamma \tilde{H}^T + V_2) = 0,$$

найдем матрицу оптимального наблюдателя (восстанавливая аргументы)

$$L_{n+1} = \tilde{\Gamma}_{n+1} \tilde{H}^T (\tilde{H} \tilde{\Gamma}_{n+1} \tilde{H}^T + V_2)^{-1}. \quad (21.70)$$

При подстановке матрицы (21.70) в уравнение (21.67) получим матрицу дисперсий ошибки оптимального наблюдателя

$$\Gamma_{n+1} = \tilde{\Gamma}_{n+1} - L_{n+1} \tilde{H} \tilde{\Gamma}_{n+1}. \quad (21.71)$$

Таким образом, соотношения (21.61), (21.62), (21.66), (21.70), (21.71) описывают алгоритм работы дискретного оптимального наблюдателя (дискретный фильтр Калмана). Эти соотношения являются рекуррентными и удобными для вычисления на ЭВМ, хотя объем расчетов получается достаточно большим.

В реальных условиях оценивание начинается с момента $t = 0$, используя начальные значения $\hat{x}_0 = \bar{x}_0$, Γ_0 и измеряя u_0 . Затем рассчитываются $\tilde{\Gamma}_1$, L_1 , Γ_1 и определяется x_1^* . При $t = h$ на основе измерения z_1 находится оценка $\hat{x}_1$ и т.д. Так как матрицы $\tilde{\Gamma}_n$, L_n , Γ_n не зависят от x_n , u_n , z_n , они при необходимости могут быть вычислены заранее и храниться в памяти ЭВМ.

Кроме оценивания по уравнениям (21.61), (21.62) достаточно часто применяется более простой наблюдатель, аналогичный по своей струк-

туре дискретному детерминированному наблюдателю (21.5) (оптимальный дискретный *предсказатель*). Для нахождения его уравнения подставим (21.62) в уравнение (21.61) и заменим $\mathbf{x}_n^*$ на $\hat{\mathbf{x}}_n$, в результате чего придем к уравнению наблюдателя, совпадающему с (21.5),

$$\hat{\mathbf{x}}_{n+1} = \Phi \hat{\mathbf{x}}_n + \tilde{\mathbf{L}}_n (\mathbf{z}_n - \mathbf{H} \hat{\mathbf{x}}_n), \quad (21.72)$$

где $\tilde{\mathbf{L}}_n = \Phi \mathbf{L}_n$. Подставив (21.71) в уравнение (21.66) с заменой $\tilde{\mathbf{L}}_n$ на $\mathbf{L}_n$, получим уравнение для дисперсионной матрицы ошибки оценивания

$$\mathbf{\Gamma}_{n+1} = \Phi \mathbf{\Gamma}_n \Phi^T - \tilde{\mathbf{L}}_n \mathbf{H} \mathbf{\Gamma}_n \Phi^T + \mathbf{G} \mathbf{V}_1 \mathbf{G}^T, \quad \mathbf{\Gamma}(0) = \mathbf{\Gamma}_0. \quad (21.73)$$

При вычислении оптимальной матрицы коэффициентов $\tilde{\mathbf{L}}_n = \Phi \mathbf{L}_n$ наблюдателя (21.72) используем выражение (21.70)

$$\tilde{\mathbf{L}}_n = \Phi \mathbf{\Gamma}_n \mathbf{H}^T (\mathbf{H} \mathbf{\Gamma}_n \mathbf{H}^T + \mathbf{V}_2)^{-1}. \quad (21.74)$$

После подстановки (21.74) в уравнение (21.73) и преобразований запишем разностное уравнение Риккати

$$\mathbf{\Gamma}_{n+1} = \Phi \left[\mathbf{I} - \mathbf{\Gamma}_n \mathbf{H}^T (\mathbf{H} \mathbf{\Gamma}_n \mathbf{H}^T + \mathbf{V}_2)^{-1} \mathbf{H} \right] \mathbf{\Gamma}_n \Phi^T + \mathbf{G} \mathbf{V}_1 \mathbf{G}^T, \quad (21.75)$$

которое является дискретным аналогом дифференциального уравнения наблюдателя в непрерывной стохастической системе (21.26) и дуальным относительно уравнения в задаче оптимального дискретного регулирования (20.71).

Стационарный оптимальный дискретный наблюдатель с постоянной матрицей $\tilde{\mathbf{L}}$ описывается уравнениями

$$\mathbf{\Gamma} = \Phi \left[\mathbf{I} - \mathbf{\Gamma} \mathbf{H}^T (\mathbf{H} \mathbf{\Gamma} \mathbf{H}^T + \mathbf{V}_2)^{-1} \mathbf{H} \right] \mathbf{\Gamma} \Phi^T + \mathbf{G} \mathbf{V}_1 \mathbf{G}^T, \quad (21.76)$$

$$\tilde{\mathbf{L}} = \Phi \mathbf{\Gamma} \mathbf{H}^T (\mathbf{H} \mathbf{\Gamma} \mathbf{H}^T + \mathbf{V}_2)^{-1}. \quad (21.77)$$

Решение алгебраического уравнения (21.76) может быть найдено как установившееся решение разностного уравнения (21.75) при $n \rightarrow \infty$ также, как в случае оптимального дискретного регулятора (20.75).

21.5.1. Стохастическая оптимальная дискретная система управления при неполном измерении состояния.

Принцип разделения

Рассмотрим задачу определения управления объектом (21.58), доставляющего минимум функционалу

$$J = M \left[\sum_{n=0}^{N-1} (\mathbf{y}_{n+1}^T \mathbf{Q}_1 \mathbf{y}_{n+1} + \mathbf{u}_n^T \mathbf{R} \mathbf{u}_n) + \mathbf{x}_N^T \mathbf{F} \mathbf{x}_N \right]. \quad (21.78)$$

Заданы вектор выходных координат объекта $y_n = Cx_n$ и вектор переменных, доступных для измерения z_n (21.59). После подстановки $y_n = Cx_n$ в (21.78) получим

$$J = M \left[\sum_{n=0}^{N-1} (x_{n+1}^T Q x_{n+1} + u_n^T R u_n) + x_N^T F x_N \right], \quad (21.79)$$

где $Q = C^T Q_1 C$.

Также как для непрерывной стохастической системы при неполном измерении вектора состояния (§ 21.4) оптимальный дискретный регулятор, доставляющий минимум критерию (21.79) для объекта (21.58) при условии (21.59), формируется на основе вектора оценки $\hat{x}$, т.е.

$$u_n = -K_n \hat{x}_n,$$

где матрица коэффициентов K_n определяется выражением (20.69), вектор оценки $\hat{x}$ вычисляется в соответствии с уравнениями оптимального дискретного наблюдателя (21.61), (21.62) или (21.72). Расчет матриц регулятора K_n и наблюдателя L_n выполняются аналогично непрерывной системе *раздельно*, независимо друг от друга. Доказательство принципа разделения для дискретной системы может быть выполнено по схеме, примененной для непрерывной системы (§ 21.4).

Применение оптимального наблюдателя приводит к увеличению критерия оптимальности (21.50) в непрерывной системе и (21.79) — в дискретной по сравнению с системой, имеющей полную и точную информацию о состоянии объекта управления. Для непрерывной системы (21.15), (21.16), (21.50) справедливы следующие выражения [39, 55]:

$\Delta J_1^H = \text{tr}(P \Gamma_0)$ — приращение функционала, обусловленное неопределенностью начального состояния $x(0)$ объекта;

$\Delta J_2^H = \int_0^{t_k} \text{tr}(K(t)^T R K(t) \Gamma(t)) dt$ — приращение критерия, вызванное ошибкой оценивания.

Аналогично для дискретной системы (21.58), (21.59), (21.79)

$$\Delta J_1^A = \text{tr}(W_0 \Gamma_0), \quad \Delta J_2^A = \sum_{i=0}^{N-1} \text{tr}(K_i^T (R + D^T W_{i+1} D) K_i \Gamma_i).$$

В случае стационарного регулятора и наблюдателя ($t_k \rightarrow \infty$, $t_0 \rightarrow -\infty$) приращения критериев (21.50) и (21.79) составят

$$\Delta J_2^H = \text{tr}(K^T R K \Gamma), \quad \Delta J_2^A = \text{tr}(K^T (R + D^T W D) K \Gamma)$$

соответственно.

Заметим, что при построении систем управления с наблюдателями возможна расхожимость оценок, обусловленных несоответствием модели объекта, характеристик шумов, начальных условий реальным данным. Оказывает влияние и погрешность вычислений на ЭВМ (ограничение числа разрядов). Для устранения расхожимости разрабатывают специальные приемы регуляризации наблюдателей [39].

Вычислительные трудности, связанные с расчетом параметров наблюдателей, с анализом процессов при работе системы управления в настоящее время преодолеваются, используя пакет прикладных программ системы MATLAB [50].

В последние десятилетия теория управления, несомненно, играла ключевую роль в развитии многих технологий. Управление — это совокупность операций по организации некоторого процесса для достижения определенных целей. Системы управления используются повсеместно: в промышленности, энергетике, транспорте, быту. Идеи и методы теории управления применяются в экономике, планировании инвестиций, процессы управления можно обнаружить в живых организмах и биологических системах.

Что же является общим для технологии управления в самых различных приложениях?

Существует прямой ответ на этот вопрос — это информация. Для управления необходима информация о цели, свойствах и состоянии объекта, воздействиях среды. Достижение основных целей управления — стабилизация неустойчивых режимов, подавление возмущений и воспроизведение задающих воздействий — связано с получением, передачей и обработкой информации. Во многих случаях именно дефицит информации ограничивает возможность управления и достижимое качество процессов.

Изучение изложенных в книге основ теории управления позволяет понять, что управление невозможно без достаточной информации, а привлечение дополнительной информации создает предпосылки для улучшения качества управления. Хотя количественно эту связь в общем виде не удастся выразить, ее можно иллюстрировать на качественных примерах. Приведем самые очевидные из них.

В системе с обратной связью действие возмущений на объект ослабляется в той мере, каково усиление контура. При конечных значениях усиления обеспечивается инвариантность до ε . Дополнительная *априорная* информация о спектре возмущения (о полюсах изображения возмущения), привлекаемая на этапе

проектирования, дает возможность синтеза алгоритма управления, обеспечивающего селективную инвариантность — нулевую установившуюся реакцию. А если удастся измерять возмущение, т.е. получать полную *текущую* информацию, то можно спроектировать компенсатор, в принципе, позволяющий добиться абсолютной инвариантности управляемой переменной к этому возмущению.

В качестве второго примера рассмотрим систему с обратной связью, регулятор которой спроектирован на базе линейной модели объекта второго порядка. Достижимое качество управления ограничивается областью адекватности модели — нельзя повышать усиление (добротность) контура на высоких частотах, где модель недостаточно хорошо описывает объект. Если же увеличить информацию о свойствах объекта, например, привлечь модели третьего и более высокого порядка (учесть малые постоянные времени и тем самым расширить частотную область адекватности модели), то можно поднять усиление контура и после этапа коррекции расширить полосу частот системы, что означает ускорение собственных составляющих процессов, способность отработки задающих воздействий и подавления возмущений более широкого спектра.

Наиболее часто по линейной модели объекта синтезируется линейный регулятор, обеспечивающий устойчивость «в малом». При больших отклонениях переменных такая система не гарантирует устойчивость положения равновесия. Если же рассмотреть нелинейную модель, адекватно описывающую объект и для больших отклонений, то соответствующими методами можно получить регулятор, обеспечивающий устойчивость положения равновесия «в большом». Ясно, что приращение качества достигнуто за счет того, что нелинейная модель более информативна. Большая информативность нелинейной модели следует и из того, что линейную модель можно получить линеаризуя нелинейную, а обратный переход неоднозначен.

Весьма важно соотношение априорной и текущей информации — чем больше первой, тем меньше требуется второй для достижения цели управления с тем же качеством процессов. Недостаток априорной информации можно дополнять путем обработки текущей информации, как это имеет место в адаптивных системах.

Число таких примеров можно увеличить многократно. Принципы управления являются результатом классификации систем именно по информационному признаку. По-существу, у

нас только два принципа управления: «по причине», когда доступна информация о возмущении (состоянии среды), и «по следствию», когда имеется текущая информация о состоянии объекта. Первый принцип разомкнутого управления позволяет компенсировать конкретное возмущение, а принцип замкнутого управления имеет универсальный характер. Обратная связь направлена на ослабление следствий — отклонений состояния объекта от цели, чем бы это отклонение ни было вызвано — предысторией системы, текущими воздействиями или изменениями свойств объекта. Обратная связь является необходимым условием стабилизации неустойчивых объектов, уменьшения влияния непосредственно неизмеряемых возмущений и уменьшения чувствительности характеристик системы к вариациям свойств объекта. Неудивительно, что в теории управления уделяется огромное внимание именно системам, реализующим принцип отрицательной обратной связи.

Дефицит информации ограничивает возможности управления. Другим обстоятельством, препятствующим достижению желаемого качества процессов, является ограниченность ресурсов управления (энергетических, материальных, финансовых). Без ограничений информационного или/и ресурсного характера нет проблемы управления.

На первый взгляд, техника управления выглядит просто и прямолинейно: компенсируй, если знаешь причины, или ослабляй их следствия с помощью обратной связи. Вместе с тем, задачи управления поражают своей утонченностью и сложностью, как в теории, так и на практике. Главным образом, это происходит из-за того, что информация запаздывает, а управляющие воздействия не могут мгновенно изменять состояние динамической системы. Даже правильные управляющие воздействия, но приложенные к объекту несвоевременно, могут вызвать катастрофические результаты. Воздействия на системы, в силу инерционности последних, изменяют их состояние не мгновенно, а во времени, и это необходимо учитывать.

В теории управления изначально принимается, что исследуемая система находится в окружении среды и взаимодействует с ней. Принятие ряда допущений позволяет обособлять объект и рассматривать искусственно разорванную цепь причинно-следственных отношений: «среда — система — среда». Анализ систем управления не ограничивается констатацией тех или иных свойств, но также объясняет, почему система с данной структурой связывает между элементами с конкретными характеристиками

имеет такие свойства и, что изменится в свойствах системы в целом, если структура системы и/или параметры того или иного элемента будут другими.

Теория управления имеет принципиально междисциплинарный характер, поскольку базируется на математических моделях, а ее результаты применимы к системам различной природы. Действительно, объекты различной природы могут иметь одно и то же математическое описание, с другой стороны, одному и тому же объекту в зависимости от целей и постановок задач управления может соответствовать множество математических моделей, различающихся параметрами, структурой и даже классом — от простейших линейных стационарных детерминированных конечномерных до нелинейных нестационарных стохастических с распределенными параметрами.

Фундаментально важно то, что занятия теорией управления позволяют понять, что общего между реальными объектами и символьными моделями. Приучают к осмотрительности в применении моделей на примерах неточного объяснения и плохого предсказания. Практические занятия дают иллюстрации того, что получается в результате анализа и синтеза систем автоматического управления, если выводы получены по математическим описаниям, недостаточно адекватным оригиналу.

Теория управления широко использует математический аппарат; даже в пределах данного учебника по основам теории были использованы определения и результаты многих разделов математики — линейной алгебры, теории дифференциальных и разностных уравнений, функционального анализа, теории функций комплексного переменного, методов оптимизации и вариационного исчисления, теории вероятностей и др.

Теория управления — наука, открытая к новым приложениям и динамично развивается в соответствии с нуждами техники и технологий, обогащая свою теоретическую и прикладную базы, привлекая новые технические средства и современные информационные технологии.

Чем бы не занимались инженеры, все более или менее значительные задачи включают моделирование, анализ, синтез и оптимизацию, т.е. включают основные компоненты, образующие сущность методологии теории управления. Умение и навыки, приобретаемые в процессе изучения теории управления, окажутся полезными в любой области деятельности.

1. Айзерман М.А. Теория автоматического регулирования.— М.: Наука, 1966.
2. Александров А.Г. Оптимальные и адаптивные системы.— М.: Высшая школа, 1989.
3. Алексеев А.А., Имаев Д.Х., Кузьмин Н.Н., Яковлев В.Б. Теория управления.— СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 1999.
4. Анализ и синтез систем управления/Имаев Д.Х., Ковальски З., Яковлев В.Б. и др.— СПб., Гданьск, Сургут, Томск: Информ. центр Сургутского гос. ун-та, 1998.
5. Андреев Ю.Н. Управление конечномерными линейными объектами.— М.: Наука, 1976.
6. Андронов А.А. Собрание трудов.— М.: Изд-во АН СССР, 1956.
7. Анисимов В.И., Вавилов А.А., Фатеев А.В. Сборник примеров и задач по теории автоматического регулирования.— М.; Л.: Госэнергоиздат, 1959.
8. Анисимов В.И. Топологический расчет электронных схем.— Л.: Энергия, 1977.
9. Арнольд В.И. Теория катастроф.— М.: Наука, 1990.
10. Атанс М., Фалб П. Оптимальное управление.— М.: Машиностроение, 1968.
11. Афанасьев В.Н., Колмановский В.Б. Математическая теория конструирования систем.— М.: Машиностроение, 1984.
12. Аязян Г.К. Расчет автоматических систем с типовыми алгоритмами регулирования/Уфимский нефт. ин-т.— Уфа, 1989.
13. Бесекерский В.А., Попов Е.П. Теория систем автоматического управления.— М.: Наука, 1975.
14. Бир Ст. Кибернетика и управление производством.— М.: Физматгиз, 1963.
15. Блэкьер О. Анализ нелинейных систем.— М.: Мир, 1969.
16. Бодэ Г. Теория цепей и проектирование усилителей с обратной связью.— М.: Иностран. лит-ра, 1948.

17. *Болтянский В.Г.* Математические методы оптимального управления.— М.: Наука, 1969.
18. *Борцов Ю.А.* Математические модели автоматических систем.— Л.: Изд-во ЛЭТИ, 1981.
19. *Вавилов А.А.* Частотные методы расчета нелинейных систем.— Л.: Энергия, 1970.
20. *Вавилов А.А.* Структурный и параметрический синтез сложных систем.— Л.: Изд-во ЛЭТИ, 1979.
21. *Вавилов А.А., Имаев Д.Х.* Машинные методы расчета систем управления.— Л.: Изд-во Ленингр. ун-та, 1981.
22. *Вавилов А.А., Имаев Д.Х.* Эволюционный синтез систем управления.— Л.: Изд-во ЛЭТИ, 1983.
23. *Вагнер Г.* Основы исследования операций.— М.: Мир, 1972.
24. *Винер Н.* Кибернетика, или управление и связь в животном и машине.— М.: Наука, 1983.
25. *Воронов А.А.* Основы теории автоматического управления.— М.; Л.: Энергия, 1965.
26. *Габасов Р.Г., Кириллова Ф.М.* Основы динамического программирования.— Минск: Изд-во Белорусского гос. ун-та, 1975.
27. *Гантмахер Ф.Р.* Теория матриц.— М.: Наука, 1967.
28. *Гарднер М.Ф., Бэрнс Дж. Л.* Переходные процессы в линейных системах.— М.: Физматгиз, 1961.
29. *Джурри Э.* Импульсные системы автоматического регулирования.— М.: Физматгиз, 1963.
30. *Дидук Г.А.* Машинные методы исследования автоматических систем.— Л.: Энергоатомиздат, 1983.
31. *Зотов Н.С.* Многоуровневый численный алгоритм расчета оптимальных по быстродействию нелинейных систем//Изв. СПбГЭТУ, вып. 482.— СПб., 1985.
32. *Иванов А.А., Фалдин Н.В.* Теория оптимальных систем автоматического управления.— М.: Наука, 1981.
33. *Изерман Р.* Цифровые системы управления.— М.: Мир, 1984.
34. *Имаев Д.Х., Краснопрошина А.А., Яковлев В.Б.* Теория автоматического управления. Ч. 1: Линейные системы автоматического управления.— Киев: Выща шк., 1992.
35. *Имаев Д.Х., Краснопрошина А.А., Яковлев В.Б.* Теория автоматического управления. Ч. 2: Нелинейные, импульсные и стохастические системы автоматического управления.— Киев: Выща шк., 1992.
36. Имитационное моделирование производственных систем/Под ред. А.А. Вавилова.— М.: Машиностроение; Берлин: Veb Verlag Technik, 1983.
37. *Катковник В.Я., Полуэктов Р.А.* Многомерные дискретные системы управления.— М.: Наука, 1966.
38. *Катханов М.Н.* Теория судовых автоматических систем.— Л.: Судостроение, 1985.
39. *Квакернаак Х., Сиван Р.* Линейные оптимальные системы управления.— М.: Мир, 1977.

40. Клюев А.С., Колесников А.А. Оптимизация автоматических систем, оптимальных по быстродействию.— М.: Энергоатомиздат, 1982.
41. Красовский А.А., Буков В.Н., Шендрик В.С. Универсальные алгоритмы оптимального управления непрерывными процессами.— М.: Наука, 1977.
42. Красовский А.А., Поспелов Г.С. Основы автоматики и технической кибернетики.— М.: Госэнергоиздат, 1962.
43. Кузовков Н.Т. Динамика систем автоматического управления.— М.: Машиностроение, 1968.
44. Куо Б. Теория и проектирование цифровых систем управления.— М.: Машиностроение, 1984.
45. Лузин Н.Н., Кузнецов П.И. К абсолютной инвариантности и инвариантности до ϵ в теории дифференциальных уравнений//Докл. АН СССР, 1946. Т. 51. N 4. С. 247—249; № 5. С. 331—333.
46. Лурье А.И. Некоторые нелинейные задачи теории автоматического регулирования.— М.: Гостехиздат, 1951.
47. Маркушевич А.И. Краткий курс теории аналитических функций.— М.: Наука, 1966.
48. Математическое моделирование/Под ред. Дж. Эндрюса и Р. Мак-Лоуна.— М.: Мир, 1979.
49. Математическая теория оптимальных процессов/Л.С. Понтрягин, В.Г. Болтянский, Р.В. Гамкрелидзе, Е.Ф. Мищенко.— М.: Наука, 1969.
50. Медведев В.С., Потемкин В.Г. Control System Toolbox/ MATLAB 5 для студентов.— М.: Изд-во ДИАЛОГ-МИФИ, 1999.
51. Моисеев Н.Н. Элементы теории оптимальных систем.— М.: Наука, 1974.
52. Наумов Б.Н. Теория нелинейных автоматических систем.— М.: Наука, 1972.
53. Нелинейные и импульсные автоматические системы/Под ред. В.Б. Яковлева.— Л.: ЛЭТИ, 1981.
54. Олейников В.А., Зотов Н.С., Пришвин А.М. Основы оптимального и экстремального управления.— М.: Высшая школа, 1969.
55. Острем К. Введение в стохастическую теорию управления.— М.: Мир, 1973.
56. Острем К., Виттенмарк Б. Системы управления с ЭВМ.— М.: Мир, 1987.
57. Павлов А.А. Синтез релейных систем, оптимальных по быстродействию.— М.: Наука, 1966.
58. Первозванский А.А. Курс теории автоматического управления.— М.: Наука, 1986.
59. Понтрягин Л.С. Обыкновенные дифференциальные уравнения.— М.: Наука, 1965.
60. Попов Е.П. Прикладная теория процессов управления в нелинейных системах.— М.: Наука, 1973.
61. Постников М.М. Устойчивые многочлены.— М.: Наука, 1981.
62. Пупков К.А., Капалин В.И., Ющенко А.С. Функциональные ряды в теории нелинейных систем.— М.: Наука, 1976.

63. *Растригин Л.А.* Системы экстремального управления.— М.: Наука, 1974.
64. Расчет автоматических систем/Под ред *А.В. Фатеева*.— М.: Высшая школа, 1973.
65. Расчет систем управления с применением СМ ЭВМ/Под ред. *В.Б. Яковлева*.— Л.: Изд-во ЛЭТИ, 1987.
66. *Ройтенберг Я.Н.* Автоматическое управление.— М.: Наука, 1978.
67. *Розенвассер Е.Н., Юсупов Р.М.* Чувствительность систем управления.— М.: Наука, 1981.
68. *Ротач В.Я.* Расчет динамики промышленных автоматических систем регулирования.— М.: Энергия, 1973.
69. *Сергеев Э.В.* Основы статистической динамики линейных систем управления.— Л.: Изд-во ЛЭТИ, 1981.
70. *Сигорский В.П.* Математический аппарат инженера.— Киев: Техніка, 1975.
71. *Табак Д., Куо Б.* Оптимальное управление и математическое программирование.— М.: Наука, 1975.
72. Теория автоматического управления/Под ред. *А.В. Нетушила*.— М.: Высшая школа, 1976.
73. Теория автоматического управления. В 2 ч. Ч. 2./Под ред. *А.А. Воронова*.— М.: Высшая школа, 1986.
74. *Терехов В.А., Ефимов Д.В., Тюкин И.Ю.* Искусственные нейронные сети и их применение в системах автоматического управления/СПб: Изд-во ГЭТУ, 1997.
75. Техническая кибернетика. Теория автоматического регулирования/Под ред. *В.В. Солодовникова*.— М.: Машиностроение, 1967.
76. *Томович Р., Вукобратович М.* Общая теория чувствительности//Сов. радио, 1972.
77. *Ту Ю.* Цифровые и импульсные системы автоматического управления.— М.: Машиностроение, 1964.
78. *Тэлор Дж., Пестель М.* Анализ и расчет нелинейных систем автоматического управления. — М.; Л.: Энергия, 1964.
79. *Уокэм М.* Линейные многомерные системы управления.— М.: Наука, 1980.
80. *Фатеев А.В.* Основы линейной теории автоматического регулирования.— М.; Л.: Госэнергоиздат, 1954.
81. *Федоров С.М., Литвинов А.П.* Автоматические системы с цифровыми управляющими машинами.— М.; Л.: Энергия, 1965.
82. *Фельдбаум А.А.* Электрические системы автоматического регулирования.— М.: Оборонгиз, 1957.
83. *Фельдбаум А.А.* Основы теории оптимальных автоматических систем.— М.: Наука, 1966.
84. Фильтрация и стохастическое управление в динамических системах/Под ред. *К.Т. Леондеса*.— М.: Мир, 1980.
85. *Флейшман Б.С.* Основы системологии.— М.: Радио и связь, 1982.
86. *Хлыпало Е.И.* Нелинейные системы автоматического регулирования.— Л.: Энергия, 1967.

87. Хофер Э., Лундерштедт Р. Численные методы оптимизации.— М.: Машиностроение, 1981.
88. Цыпкин Я.З. Основы теории автоматических систем.— М.: Наука, 1977.
89. Черноусько Ф.Л. Оценивание фазового состояния динамических систем.— М.: Наука, 1988.
90. Честнат Г., Майер Р. Проектирование и расчет следящих систем и систем регулирования.— М.: Госэнергоиздат, 1959.
91. Чураков Е.П. Оптимальные и адаптивные системы.— М.: Энергоатомиздат, 1987.
92. Щипанов Г.В. Теория и методы проектирования автоматических регуляторов//Автоматика и телемеханика N 1. С. 49—66. 1939.
93. Bernstein D.S. A Student Guide to Classical Control//IEEE Control Systems Magazine. 1997. N 4. P. 96—100.
94. Demuth H., Beale M. Neural Network Toolbox for Use with MATLAB.— The MathWorks, Inc., Natick, MA, 1994.
95. Demuth D., Hellendoorn H, Reinfrank M. Introduction to Fuzzy Control.— Berlin, Heidelberg: Springer Verlag, 1993.
96. Gulley N., Jang R.J.-S. Fuzzy Logic Toolbox for Use with MATLAB.— The MathWorks, Inc., Natick, MA, 1995.
97. Kosko B. Neural Networks and Fuzzy Systems.— Englewood Cliffs, NJ; Prentice Hall, 1992.
98. Samad T. Special Issue on Emerging Technologies; Guest Editor's Introduction//IEEE Control Systems Magazine. 1997. N 6. P. 16—18.
99. Silver E.A., Peterson R. Decision Systems for Inventory Management and Production Planning.— John Wiley&Sons, Inc., 1985.

**Душин Сергей Евгеньевич, Зотов Николай Сергеевич,
Имаев Дамир Хабибович, Кузьмин Николай Николаевич,
Яковлев Владимир Борисович**

ТЕОРИЯ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Редакторы Е.Н. Рожкова, Л.В. Честная

Художник А.Г. Игнатьев

Художественный редактор З.Е. Анфиногенова

Технические редакторы Н.В. Быкова, Л.А. Овчинникова

Корректоры В.А. Жилкина, О.Н. Шебишова

Лицензия ИД № 06236 от 09.11.01.

Изд. № ВТИ-158. Сдано в набор 16.09.02. Подп. в печать 16.06.03.

Формат 60×88¹/₁₆. Бум. газетная. Гарнитура «Таймс». Печать офсетная.

Объем: 34,79 усл. печ. л., 35,29 усл. кр.-отт.

Тираж 3000 экз. Заказ № Я-440

**ФГУП «Издательство «Высшая школа», 127994, Москва, ГСП-4,
Неглинная ул., 29/14.**

Тел.: (095) 200-04-56

E-mail: info@v-shkola.ru <http://www.v-shkola.ru>

Отдел реализации: (095) 200-07-69, 200-59-39, факс: (095) 200-03-01.

E-mail: sales@v-shkola.ru

Отдел «Книга-почтой»: (095) 200-33-36. E-mail: bookpost@v-shkola.ru

Отпечатано в ГУП ПИК «Идел-Пресс».

420066. г. Казань, ул. Декабристов, 2.